SINGLE NEURON ANALYSIS AND COMBINATORIAL FUNCTION OF DELTA-PROTOCADHERINS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

By

Adam Bisogni

August 2017

SINGLE NEURON ANALYSIS AND COMBINATORIAL FUNCTION OF DELTA-PROTOCADHERINS

Adam Bisogni, Ph.D.

Cornell University 2017

The average adult human brain consists of approximately 86 billion neurons and roughly 14 x 10^{15} synapses. How this marvelously complex system is achieved with great precision throughout development is largely unknown. To date, roughly 200 genes have been identified that play direct roles in aspects of neural wiring. How can such a small number of genes specify such a vast amount of instructions for neural wiring? It is clear that individual neurons must express these genes in combinations to greatly expand their utility towards specifying neural wiring to such great depths of complexity. However, two major roadblocks have prevented an understanding of how this occurs.

First, the nature and extent of these combinations have not been defined, and so the basic aspects of combinatorial expression such as the depth and range of possible combinations are unknown. Improving our understanding in this respect will allow us to begin addressing the second limitation, which is understanding what the combinations functionally achieve for any given neuron. To address the first limitation, new approaches that can profile gene expression in single cells can be utilized to define the combinations. To address the second limitation regarding function, new complementary approaches need to be developed. For example, to study how a gene family works together, first the function of its individual members on neuronal behavior need to be defined. Next, the effect of expressing multiple family members needs to be tested, until

an understanding of the combinatorial functions within this gene family becomes apparent. Finally, the combinations of this gene family could be expanded and integrated with different genes or gene families to a broader understanding of combinatorial function.

The aim of this thesis is to explore this concept by studying one small gene family of cell adhesion molecules, the δ-protocadherins. In Chapter 1, an overview of the complexities of neural wiring, the types of molecular cues that provide instructions, and the importance of understanding their functions in combinations are discussed. In Chapter 2, a single-cell gene expression analysis method is developed and used to define the extent to which the δ-protocadherins are expressed in combination in single olfactory sensory neurons. In Chapter 3, the consequences of combinatorial δ-protocadherin expression on their adhesive functions are studied, and a simple model is proposed that explains combinatorial adhesive specificities based on differential apparent affinities, and the relative proportion of surface expression of each gene. Finally, in Chapter 4, I explore how the model could be extended and integrated with other protocadherins and to create a larger model that could encompass the entire cadherin superfamily.

BIOGRAPHICAL SKETCH

Adam Joseph Bisogni was born in Ithaca, New York in 1986. He graduated from Ithaca High School in 2004. In 2008, he received a bachelor's degree with honors in Human Biology, Health, and Society from Cornell University. In 2010, he started his Ph.D. in the Department of Biomedical Sciences at Cornell University under the supervision of Dr. David Lin. He completed his Ph.D. work on δ-protocadherins in 2017.

Dedicated to Mom, with all my love.

ACKNOWLEDGEMENTS

I would like to first thank my advisor, Dr. David Lin. His mentorship style extended my love for science. He always makes himself available to discuss science and experiments even when my ideas are crazy. You can tell he loves science by the fact that he spends just as much time at the bench than at his office desk. I am thankful for his active participation in lab, and willingness to support any experiment that may be going on in the lab.

I would also like to thank the other members of my committee, Dr. John Schimenti, Dr. David Deitcher, and Dr. Chris Schaffer for their input and feedback over the years.

It was a pleasure collaborating with Dr. Jean Yang and Dr. Shila Ghazanfar in the School of Mathematics and Statistics at the University of Sydney, Australia. I learned a lot from our discussions, and I thank them greatly for their contributions on the statistical analyses of the single-cell data.

Finally, I would like to thank my family and friends. You all supported me in so many ways. Special thanks to Catalina. And most importantly, my parents, for always encouraging and supporting my education and goals. I hope I've made you proud by becoming the third Cornell Ph.D. in the family.

Table of Contents

Chapter 1

Introduction

1. The Complexities of Brain Development

Since the birth of modern neuroscience over 100 years ago (Cajal 1909), a major goal has been to understand how brains are constructed. Answering this question has profound implications. A complete understanding of mammalian brain development would have an immense impact on science, medicine, and humanity. Obtaining this knowledge, however, is an extremely daunting task due to the enormous level of dynamic complexity that exists within a developing brain. This complexity is best illustrated by considering the magnitude of change that occurs between the birth of the first neurons during embryonic development, and adulthood. For example, in the mouse, the first clusters of cells to express the neuronal marker beta-tubulin III appear at embryonic day 8.5 in the neural plate (Easter et al. 1993). By the time the mouse reaches adulthood several weeks later, there are approximately 71 million neurons in the brain (Herculano-Houzel et al. 2006) that have formed what can be roughly estimated to be 6.5 trillion synapses. Even greater complex dynamism occurs in humans. The first neurons are produced at embryonic day 42 (Stiles and Jernigan 2010), and by adulthood there are approximately 86 billion neurons in the brain (Azevedo et al. 2009). Staggeringly, the number of synapses in the adult human can be estimated to be on the order of 14 quadrillion (14×10^{15}). Thus, between the "birth" of

the first neurons and adulthood, a remarkable series of concerted molecular and cellular events organizes a massive number of neurons into an even greater number of organized synapses. These processes establish the rules for ordering neural circuit architecture, allowing for the creation of specialized networks, which ultimately allows for the emergence of both simple and complex behaviors. Immense efforts have been made to identify genes and understand their function in neural development. However, despite decades of work the function of many genes remain poorly understood, and even less is known about how these genes work together to regulate development. Addressing this issue is pivotal in order to untangle the depth of complexity which exists in neural development. Only by beginning to study how these processes work together will we eventually be able to progress towards a fundamental understanding of brain development, and better understand brain function in both health and disease.

2. The Molecular and Cellular Orchestration of Brain Development

2.1. Evidence for a Genetic Blueprint

What is required to orchestrate and instruct the migration, guidance, and wiring of neurons to ultimately form a mature brain that can correctly transmit, receive and process information? Despite being the most complex organ with extremely high degrees of both cellular and molecular heterogeneity (MacNeil and Masland 1998; Usoskin et al. 2015; Zeisel et al. 2015; Darmanis et al. 2015; Tasic et al. 2016; La Manno et al. 2016), the brain clearly has evolutionarily conserved programs in place

that offers a blueprint for development. Defining what these programs are provides direction for studying their functional impact in order to build a framework towards understanding neural development. Examples supporting the existence of a genetic blueprint can be observed at various levels of analysis. On a transcriptional level, several studies have used microarray profiling to examine thousands of anatomically defined regions of human and rhesus macaque post-mortem brain tissues to understand the dynamics of the transcriptional landscapes in the developing brain. Strikingly, they all found that only a small amount a variation is observed when comparing the same regions across different individuals at similar stages of development. In contrast, high levels of variation are observed when comparing profiles across different regions, or developmental stages from the same individual (Colantuoni et al. 2011; Kang et al. 2011; Hawrylycz et al. 2012; Bakken et al. 2015). Thus, despite differences in sex, millions of sequence differences across genomes (Jorde and Wooding 2004; Levy et al. 2007), or the innumerable unique environmental events experienced by each individual, the transcriptional landscape of healthy "normal" brains appears to be remarkably consistent in a spatial and temporal manner. This suggests that a genetic blueprint establishes a significant set of foundational rules for programming brain development.

On a functional level of analysis, many genes involved in neural wiring have been found to be highly conserved across species (Dickson 2002). An example of a highly conserved mechanism is the commissural midline choice, where functional studies have revealed similar genetic mechanisms for both vertebrates and some invertebrates

(Kennedy et al. 1994; Serafini et al. 1994) to instruct peripheral commissural neurons to cross the midline of the body plan. From worm to humans, this particular mechanism has been conserved over more than 600 million years of evolution (Chisholm and Tessier-Lavigne 1999). However, a smaller number genes have also been identified to be divergent and not conserved. In such cases the genes either lack homologs, or have homologs with unknown or non-homologous function. For example, VAB-8 is critical for cell migration and axon guidance in *C. elegans* (Manser and Wood 1990; Wightman et al. 1996; Wolf et al. 1998; Watari-Goshima et al. 2007), but has no identified homologs in other species.

Finally, gross anatomy of the brain can also highlight the consequences of divergent evolution and the result of different wiring programs. Even within the same evolutionary class, differences in major axonal tracts can be observed. The most striking example is the corpus callosum, which is a large interhemispheric tract of axons that enhances neural communication between the two cerebral hemispheres (Gazzaniga 2000; Phillips et al. 2015). It is only found in the Eutherian (placental) mammals, and is noticeably absent from the Metatheria (marsupial) and Prototheria (monotreme) subgroups (Owen 1837; Flower 1865; Smith 1910; Katz et al. 1983). Additionally, the basic fiber characteristics of the corpus callosum in placental mammals has been found to be well conserved across animals of the same species and across different species (Aboitiz et al. 1992; Olivares et al. 2001; Phillips et al. 2015), suggesting strong genetic control over its development. Intriguingly, many classes of genes involved in wiring appear to be conserved across members of the three classes (personal observation), but having

diverged from a common ancestor roughly 200 million years ago, exhibit significantly different genomic organization (Graves 1996; Deakin et al. 2012). This raises an interesting possibility about how differences in genomic organization may create functionally different wiring programs even when provided with similar overall genomic content. Therefore, although there is clearly a broad genetic blueprint that is widely conserved, aspects of it can vary, even among subgroups.

2.2. Modulation of the Genetic Blueprint via Activity

If such a strong blueprint exists for neural development, then how do individuals become unique? It has become clear that activity (sensory and spontaneous) is critical for remodeling the nervous system, both during development and during learning to best adapt to the challenges of the environment. For example, experiences and activity can provide the input for modulating synaptic strengths (Bruner and Tauc 1966; Castellucci et al. 1970; Trachtenberg et al. 2002), which are thought to be the fundamental storage unit for learning and memory (Grutzendler et al. 2002; Reijmers et al. 2007).

The influence of experience is most dramatically seen during synaptic pruning, where a dramatic decrease in synaptic density in humans (Huttenlocher 1979) and primates (Bourgeois et al. 1994) occurs in response to experiences after the first few years of life. Similarly, certain aspects of neural development have been found to be required to occur in specific "critical periods" to ensure proper development. This is best exemplified through the series of the classic experiments by Hubel and Wiesel, who

found that temporary monocular vision in cats during a critical time in development irreversibly disrupts ocular domain column maturation and vision in the eye that was temporarily closed. (Hubel and Wiesel 1963; Wiesel and Hubel 1963a; Wiesel and Hubel 1963b).

On a molecular level, numerous genetic studies have uncovered several genes involved in neural wiring that are regulated in an activity dependent manner in certain neural systems. Examples of such processes include axon guidance (Serizawa et al. 2006; Williams et al. 2011), dendritic growth (Nedivi et al. 1998), dendritic branching (Ohnami et al. 2008; Hayano et al. 2014) and both synapse formation and elimination (Corriveau et al. 1999; Hermey et al. 2013). Thus, despite the appearance of an evolutionary conserved, hardwired developmental program, it can and actually needs to be modulated to various degrees by activity for proper development. Combined, these two systems establish boundaries and basic rules while permitting some randomness or change (via experiential, sensory, or spontaneous neural activity) to influence neural wiring. Identifying and understanding activity regulated genes is therefore critical because they could act as functional modulators to other genes when expressed in combinations.

3. Requirements of Neural Wiring, Genetic Limitations, and Combinatorial Solutions

The power of neurons comes from their ability self-organize, creating networks and circuits. This organization creates the capacity to communicate and influence each

other, generating vast possibilities for complex activity patterns between them. Thus, *how* circuits are built becomes crucial, because the circuit formation defines the functional possibilities of that circuit. Ultimately, the collective repertoire of an animal's neural circuitry defines the extent of its behavioral possibilities. The two processes that define neural circuit assembly are axon guidance and synaptogenesis.

Before a neuron can form a synapse with another neuron, it must first extend its axon through the cellular milieu to come into proximity with the dendritic processes of other neurons. Critically, the axon must be instructed to avoid or reject certain cellular targets, while selectively being attracted towards others. This pathfinding process is mediated by various factors called axon guidance molecules, which act upon the growth cone of the axon and influence the direction of extension. This process often proceeds in a step-wise manner, with neurons finding intermediate guidepost targets or cues before reaching the final target (Grenningloh et al. 1991; Kehayova et al. 2011). In many cases, initial neurons act as 'pioneers', and future neurons are dependent upon their successful targeting to also reach their targets, although to various degrees depending on the species (Harrison 1910; Bate 1976; Bentley and Caudy 1983; Lin et al. 1995).

Mechanistically, guidance is physically achieved through signaling cascades which alter focal adhesions (Myers et al. 2011) and/or remodel the cytoskeleton of the growth cone (Letourneau et al. 1987; Lin and Forscher 1995; Challacombe et al. 1996; Lee and Suter 2008) to promote or reduce motility in certain directions. Finally, once the axon has reached the appropriate target, it must form a synapse in order to establish a connection that allows for neural transmission to occur between the two neurons. This is

a particularly dynamic process in which the synapse can be strengthened, weakened, or removed over time. Since the fundamental units of the nervous system are synapses, it is no surprise that many neurological diseases are caused by, or linked to, genes involved in axon guidance (Engle 2010; Nugent et al. 2012; Van Battum MSc et al. 2015) and synaptogenesis (Lepeta et al. 2016; Fukata and Fukata 2017).

How many genes are required to mediate these processes? To date, roughly 200 axon guidance and synaptogenesis molecules have been identified. Not considering multiple splice isoforms, this means that roughly only 1% of the genome contributes towards wiring the billions of neurons and their trillions of synapses in the brain. How can such a small number of genes generate enough specificities to precisely wire the brain? The work of Roger Sperry in the 1940s began to address this issue. After publishing several formative experiments on the ability of surgically scrambled peripheral nerves to regenerate and regrow in a functional manner without any "re-educative adjustments" (Sperry 1943a; Sperry 1943b; Sperry 1945; Sperry 1948), he outlined his seminal "chemoaffinity hypothesis" (Sperry 1963). In this paper, he posited that the only way neurons could grow properly would be if they had "some kind of individual identification tags" that were "cytochemical in nature" and would allow for neurons to differentiate themselves between targets and other neurons by differences in selective affinities (Sperry 1963). He states that such a scheme "…requires millions, and possibly billions, of chemically differentiated neuron types, each distinguishable from all others…", but how can so many tags be achieved with only ~200 different genes? Although he did not explicitly offer a solution in his paper, the only way to create such a vast number with a

limited pool of genes is use the cytochemical tags in a combinatorial manner within individual neurons to generate a sufficient diversity of cell surfaces. While it is clear that axons and growth cones express numerous mRNAs (Zivraj et al. 2010; Gumy et al. 2011), most studies only seek to understand the function of one to two genes at once leaving the outcome of combinatorial functions unknown. Today, over 50 years since proposed in full, Sperry's chemoaffinity hypothesis remains widely accepted. However, experiments to support it and begin to understand how it might work have not progressed. New approaches are needed to reveal the molecular mechanisms behind Sperry's hypothesis. Doing so would make a significant impact in our understanding of neural wiring and brain development.

4. Identification of Axon Guidance Cues

Over 125 years ago, Ramon y Cajal discovered the growth cone of the axon by studying the direction of dorsal commissural axons and ventral motor neurons in the spinal cord of the developing chick embryo in fixed histological stains (Ramón y Cajal 1890). He routinely noticed that the growth cone of these neurons followed the same path "…without deviations or errors, as if guided by an intelligent force" (Sotelo 2002). This, along with other observations, led him to develop his "neurotropic hypothesis" (Ramón y Cajal 1892) (a clear precursor to Sperry's "chemoaffinity hypothesis") and argued that neurotropic factors elicited certain behaviors from growth cones, by providing them instructions on where to migrate. Although it was difficult to provide experimental evidence in embryos at the time, Cajal was able to perform regeneration experiments

with grafts of peripheral nerves that supported his hypothesis that the instructive cues originated from living cells, as grafts that had been treated with chloroform failed to guide regenerating axons (Ramón y Cajal 1928).

Almost 100 years passed before further evidence emerged to provide additional support for the idea behind the neurotropic hypothesis. Lumsden and Davies used collagen gel matrices to co-culture embryonic mouse sensory neurons and their target tissue at a stage prior to when the target would be reached by the neurons (Lumsden and Davies 1983). They found that neurites grew selectively towards their specific target tissue, and never towards non-target tissues. Targeting still occurred even when blocking antibodies to nerve growth factor (NGF) were added, suggesting that some form of specific chemotropism occurs to attract developing neurons to their targets (Lumsden and Davies 1983). The first experimental evidence for a diffusible, attractive guidance molecule was demonstrated in 1988 by showing that the ventral floor plate of the rat selectively attracted outgrowth of dorsal commissural neurons (Tessier-Lavigne et al. 1988). A few years later, the secreted factor was identified as netrin-1 and netrin-2 (Serafini et al. 1994), the products of genes which are homologous to the previously identified *C. elegans* gene UNC-6 (Hedgecock et al. 1990; Ishii et al. 1992). Soon after the receptor that mediates its response, DCC, was also discovered (Keino-Masu et al. 1996). A recent paper has provided an updated view that the critical source of netrin-1 may be in the ventricular zone and that it may not act as a long range diffusible cue, although it still remains important for the commissural midline choice (Dominici et al. 2017).

A few years later repulsive cues were discovered. Using the embryonic rat olfactory system, Pini co-cultured the septum and olfactory bulb explants using collagen gel matrices. He showed that a diffusible repulsive signal in the septum caused the axons of mitral and tufted cells in the olfactory bulb to project their axons away from their origin near the septum and instead migrate laterally, consistent with their lateral projection in vivo (Pini 1993). That same year, a protein called Collapsin (now Sema3A) was found to cause repulsion via growth cone collapse (Luo et al. 1993).

The further advancement of molecular and biochemical techniques led to the identification of several more conserved axon guidance molecules and their functions to either act as attractive or repulsive cues. These include additional members and receptors of the Netrin family (Rajasekharan and Kennedy 2009), the Slits and Robos (Brose et al. 1999; Kidd et al. 1999; Yu et al. 2014), Ephrins and the Eph receptors (Poliakov et al. 2004), Plexins (Tamagnone et al. 1999), Neuropilins (Parker et al. 2012; Guo and Vander Kooi 2015), the Semaphorins (Kolodkin et al. 1992; Luo et al. 1993; Fan and Raper 1995; Kobayashi et al. 1997) as well as non-canonical cues such as Ryk (receptor related to tyrosine kinases) (Hovens et al. 1992; Bonkowsky et al. 1999; Yoshikawa et al. 2001; Lu et al. 2004; Keeble et al. 2006). The range of cues becomes even more diverse when other factors that have also been found to influence guidance are also considered, such as extracellular matrix proteins (Faissner and Kruse 1990; Neugebauer et al. 1991; García-Alonso et al. 1996), cell adhesion molecules (Matsunaga et al. 1988; Neugebauer et al. 1988; Bixby and Zhang 1990; Riehl et al. 1996; Schmucker et al. 2000), growth factors (Gundersen and Barrett 1979; Gundersen

and Barrett 1980; Caton et al. 2000; Genç et al. 2004), chemokines (Xiang et al. 2002; Chalasani et al. 2003; Lieberam et al. 2005), and morphogens (Trousse et al. 2001; Liu et al. 2003; Yamauchi et al. 2008; Gordon et al. 2010). Therefore, it can be seen that a wide range of molecules, each with different functional possibilities, contribute together towards axon guidance.

5. Synaptogenesis

Once an axon has reached its final target region, the final step in the wiring of the two neurons is synaptogenesis. The growth cone must differentiate into axonal terminals and prepare for the possibility of synapsing with dendrites of other neurons. Synapses are the fundamental unit of transmission, processing, and storage of neural information (Mayford et al. 2012). For proper synapse formation to occur, a series of steps must occur on both the pre- and post-synaptic neurons to assemble a stable connection that allows neural transmission to occur between the two neurons (Friedman et al. 2000; Zhai et al. 2001; Okabe et al. 2001; Washbourne et al. 2002). The number of proteins localized to post-synaptic densities has been found to be in the hundreds (Yoshimura et al. 2004).

After reaching the target, the pre-synaptic neuron must face another question of specificity - which neurons and synapses should it synapse with? How is this choice specified? Some information is known regarding how neurons may avoid forming synapses with some neurons, and how excitatory and inhibitory synapses are specified through differentiation once formed (Song et al. 1999; Ben Chih et al. 2005; Levinson et

al. 2005). However, little is known about how synapse formation beyond that is specified. Several features of synapses make answering this question difficult. One aspect is the magnitude of synapses that the average neuron has, estimated to be in the tens of thousands (Pakkenberg 2003; Pelvig et al. 2008). It is also known that within a neuron, localized and spatially restricted mRNA transport and protein synthesis within axons and synapses occurs (Koenig 1967; Giuditta et al. 1968; Steward and Levy 1982; Steward and Ribak 1986; Greenough et al. 2001; Ostroff et al. 2002; Kanai et al. 2004) and these could potentially enable each synapse to be differentially labeled with combinations of surface proteins. Therefore, sub-dendrite/axon analysis may be required to fully understand synaptic diversity within single neurons before specificities can fully be understood. Finally, the stability and lifespan of synapses are plastic and dynamic. They can be stable for periods of weeks to months, or only a few hours to days (Trachtenberg et al. 2002). Thus, because they are able to turn over in such a dynamic fashion, it is difficult to study them as "static" processes.

Critical to creating specific synapses are trans-synaptic proteins. Typically, these are cell-adhesion proteins, and several have been localized to synapses and/or shown to play a role in synapse formation or organization. Some appear to be generic synaptic proteins that may not confer a high degree of synaptic specificity, such as Synaptic Cell Adhesion Molecule (SynCAM) (Biederer et al. 2002; Fowler et al. 2017). In contrast, members of gene families have been found to be differentially expressed in neurons and localized to synapses such as the cadherins (Togashi et al. 2002) and the integrins (Einheber et al. 1996; Nishimura et al. 1998; Pinkstaff et al. 1999; Mortillo et al. 2012).

One model proposes that proteins conferring high specificity initiate the formation of a synapse by bringing the pre- and post-synapses together in proximity, and then other synaptic cell adhesion molecules such as SynCam and Neuroligins reinforce the synapse and specify if it will become excitatory or inhibitory. Further studying how these many proteins function in concert at the synapse will greatly enhance our understanding for how neural wiring is specified.

6. A Continuous Transition from Axon Guidance to Synaptogenesis

Historically, axon guidance and synaptogenesis were considered similar but distinct processes in neural development. However, since synaptogenesis begins to occur when the pathfinding growth cone differentiates into axon terminals, a more realistic model may be to place axon guidance and synaptogenesis on the same functional continuum rather than as separate or distinct processes. This idea is supported by the findings that several axon guidance genes have been found to also play roles in synapse formation.

For example, in *C. elegans*, synapse formation between two types of interneurons was found to be mediated by glial cells acting as guideposts that express UNC-6 (netrin). In the post-synaptic neuron, the netrin receptor UNC-40 (DCC) is expressed and functions conventionally to attract the axons towards the glial cells. However, in the pre-synaptic neuron, UNC-40 plays an unconventional role, increasing the formation of pre-synaptic terminals to support synaptogenesis with the post-synaptic neuron (Colón-Ramos et al.

2007). In contrast, UNC-6 has also shown to prevent synapse formation via its UNC-5 receptor in specific motor neurons in *C. elegans* (Poon et al. 2008).

More examples of axon guidance molecules that can also regulate synaptic function can be found in the family of Eph receptors, which have been shown to regulate various aspects of synapse formation and function. These include influencing NDMA receptor clustering (Dalva et al. 2000), dendritic motility (Kayser et al. 2008) dendritic morphology (Murai et al. 2003; Henkemeyer et al. 2003), synapse maturation (Aoto et al. 2007; Lim et al. 2008), synaptic transmission (Essmann et al. 2008), and long-term potentiation (Grunwald et al. 2001; Contractor et al. 2002). Similar functions at the synapse have also been found with the Semaphorins (Sahay et al. 2005; Burkhardt et al. 2005; Morita 2006; Yamashita et al. 2007), Slit/Robos (Campbell et al. 2007; Essmann et al. 2008) and Netrin/DCC (Friedman et al. 2000). Fewer cases of synaptic proteins influencing axon guidance have been revealed, although it has been shown to occur, such as the ability of SynCAM to negatively regulate the growth cone (Stagi et al. 2010).

Taken together, these examples show the importance of considering genes in neural wiring to have the potential to be both axon guidance molecules and synaptogenic molecules. Additionally, it is critical to consider that if a gene has a defined function in axon guidance, it does not necessarily mean it will exhibit the same affect in synapse formation. Instead, it may mediate different or unique affects depending on what other interacting proteins are available in proximity.

7. Cell Adhesion

Cell adhesion molecules are a major category of genes that facilitate aspects of both axon guidance and synaptogenesis. Critical to most aspects of multicellular life, cell adhesion molecules allow cells to assemble into larger groups of cells. Importantly, through differential adhesion, cells can be selectively arranged within groups of cells, or separated entirely from others. This allows organisms to develop complex tissue and cellular organizations, something that becomes more important as organisms become larger and more complex. This is true nowhere more than in the nervous system, where such cellular distinctions become imperative in specifying neural circuit construction. But how exactly how can cell adhesion molecules accomplish this in the nervous system, and if this strategy of differential adhesion is employed, how many different cell adhesion molecules are sufficient to selectively wire the brain?

7.1. Early Experiments of Selective Adhesion

At the turn of the 20th century, the first clues for how multicellular organisms develop came from Paul Wilson's now famous experiments on sponges, the simplest extant multicellular organisms that are known to exist. He first discovered that sponges could be dissociated into single cells, and a portion of the cells will recover and begin re-aggregating into multicellular forms (Wilson 1907). Since some sponges have different pigmentations, he was then able to test whether cells from different species that were dissociated and then allowed to intermingle while recovering would re-aggregate indiscriminately, or in a selective species specific manner. In all cases, he found that

"…the cells and cell-masses of a species combined, and not the cells of different species." (Wilson 1907). This seemingly simple test is considered the first experimental demonstration of selective self-recognition in cells. Follow up studies confirmed this phenomenon in other sponges (Galtsoff 1923) and set the stage for future work in this nascent field to determine the factors that enable self-organization to produce specific tissues arrangements.

Nearly 50 years later, the principle of self-recognition between different cells was greatly expanded by Townes and Holtfreter in a landmark paper (Townes and Holtfreter 1955). Using amphibian embryos that had undergone neurulation, germ layers were separated by dissection, and then dissociated into single cells by treating them with an alkaline solution. If the pH was returned to normal, the cells would begin to re-aggregate. This allowed them to test whether cells from different germ layers would re-aggregate into distinct populations. Although initially they re-aggregate indiscriminately, over time they always spatially sorted out - usually with one cell type enveloping the other. Furthermore, any test of two or more cell types resulted in spatial arrangements that reflected the embryo with regards to one population spreading over, or invaginating the other. They hypothesized that their observations were the result of selective intercellular adhesion, and the tendency for some cells to undergo directed cell movements. They also posited that these relationships should be able to change over the course of development, as the cells within each layer further differentiate and require more specific and nuanced relationships with other cell types. This watershed paper marked

the first experimental evidence that different cell types are rearranged in the embryo based on selective affinities (Townes and Holtfreter 1955).

For a period of several years following this paper, various lab groups sought to uncover the responsible agents for endowing cells with selective affinities, and how such agents might control the phenomenon of cell sorting (Moscona and Moscona 1952; Trinkaus and Groves 1955; Moscona 1957; Steinberg 1958; Curtis 1961; Steinberg 1963; Roth and Weston 1967; Takeichi 1977; Urushihara et al. 1979). The result of this collective momentum led to the identification of the first two cell adhesion molecules to be identified: Neural-Cell Adhesion Molecule (NCAM) (Brackenbury et al. 1977; Thiery et al. 1977; Rutishauser et al. 1978a; Rutishauser et al. 1978b), and the calcium-dependent E-Cadherin (Urushihara and Takeichi 1980), providing the ability to begin studying mechanisms of how selective adhesion works.

7.2. The Cadherin Superfamily

Since then, a rich diversity of cell adhesion molecules has been identified that mediate a wide range of biological phenomenon across all multicellular life forms, particularly in nervous system function and development. Of these, the largest and possibly most important gene family is the Cadherin superfamily of calcium-dependent cell adhesion molecules, which are classified based on their IgG-like extracellular domains (Takeichi et al. 1988) Comprised of approximately 114 members in humans, and 119 in the mouse (Hulpiau et al. 2016), their biological functions are widespread. For example, outside the nervous system, they have been implicated in processes that can range

from inner call mass formation (Shirayoshi et al. 1983), to somite (Radice et al. 1997; Linask et al. 1998; Horikawa et al. 1999) and limb bud formation (Yajima et al. 1999; Omi et al. 2002) and the localization and motility of oocytes in *Drosophila* (Godt and Tepass 1998; Niewiadomska et al. 1999).

Critically, these processes rely on cadherins to provide the cells with selective and differential adhesive properties. Cells that express different cadherins will have different adhesive specificities, and will prefer the formation of bonds with cells that express the same cadherin (homophilic binding) rather than a different cadherin (heterophilic binding). A formalized model of this was proposed in 1960s by Malcolm Steinberg. Aptly named the Differential Adhesion Hypothesis (DAH) (Steinberg 2007), it posits that the behavior of cell sorting can be explained through the physical analogy of surface tension and immiscible liquids. In this model, cellular rearrangements are caused by weaker bonds being replaced by stronger bonds, until interfacial free energies are sufficiently minimized. Although debate lingered over the causative property or properties of the DAH observations (Harris 1976), the principle, theory, and predictive applications behind the DAH are abundantly supported through experimental evidence and therefore remains widely accepted today (M S Steinberg and D R Garrod 1975; Steinberg and Takeichi 1994; Duguay et al. 2003; Foty and Steinberg 2005; Schötz et al. 2008).

The DAH provides a potential mechanism that neurons could use to confer specificity for potentially each aspect of neural wiring to produce complex organizations of neurons and synapses. For example, this phenomenon could be used to mediate specific

neuronal migration, aspects of fasciculation, or the proper matching of synaptic partners. Indeed, there is a large body of evidence emphasizing the importance of cadherins across all components of neural development, such as neurulation (Hatta and Takeichi 1986), neural migration, establishment of cell specific layers, boundaries, and nuclei (Szabó et al. 2015), neurite extension, axon guidance and fasciculation (Price et al. 2002; Schwabe et al. 2014), and target recognition and synapse function (Togashi et al. 2002; Duan et al. 2014). However, the cadherins that have been the most studied are the ~20 type I "classical" cadherins and the type II cadherins. Even though cells clearly express them in combinations (Matsuyoshi and Imamura 1997; Shan et al. 2000), there are not enough of them and they are also expressed too broadly to be useful for creating adhesive specificity codes. This lead researchers to begin looking at the remaining members of the cadherin superfamily.

7.3. The Protocadherins

The largest sub-family within the cadherin superfamily are the protocadherins. The first members of the protocadherin family were discovered in 1993 using degenerate PCR targeting the IgG-like extracellular domains of classical cadherins with rat and human cDNA (Sano et al. 1993). Because fragments of the sequences identified were also found in lower organisms such as *Xenopus*, *Drosophila*, and *C. elegans*, Sano et al. considered these sequences to contain a "primordial" cadherin motif and thus coined them protocadherins. Soon after, many more protocadherins were discovered, and since then nearly 80 different protocadherins have been identified in mammals (Jontes 2013; Hulpiau et al. 2016).

Although they share similarities to the classical cadherins, their differences are of significant consequence. For example, classical cadherins have five extracellular (EC) domains, where the protocadherins have six, seven, or even up to ~40 EC domains. Furthermore, they do not bind to β-catenin as the classical cadherins do, and instead have highly variable intracellular cytoplasmic domain (ICD) sequences, suggesting a multitude of possible signaling mechanisms. Such variable characteristics have enabled the ~80 protocadherins to be further subdivided into protocadherin sub-families (Nollet et al. 2000).

The evolutionary history of protocadherins suggest that the family greatly expanded during the metazoan explosion (Hulpiau and Van Roy 2011), leading some to suggest that their emergence enabled the mammalian brain to expand and become more complex. In expression studies, their dynamic, strong, and punctate patterns throughout the mammalian brain and nervous system made them excellent candidates for conferring neurons with diversity for wiring through possible combinatorial expression patterns.

However, since their initial discovery, the protocadherins have been shown to be richly complex in both their expression and function and as a consequence, not at all straightforward to understand. While some protocadherins have been extensively studied and show critical roles in neural wiring, many remain poorly understood, or barely studied. However, their importance is neural circuit formation is highlighted by their causative or associated nature with many different neurological conditions. In addition, like the cadherins, they also play critical roles in tissues outside of the nervous

system. Here, the two largest subfamilies of the protocadherins (clustered protocadherins and non-clustered protocadherins) and their roles in neural circuit formation are discussed.

7.3.1. Clustered Protocadherins: Genomic Organization and Regulation

The clustered protocadherins are the largest sub-family of the protocadherins, consisting of 60 genes (Hulpiau et al. 2016). A unique characteristic shared among the first members to be identified set the stage for uncover their striking genomic organization. Within the first two groups discovered, all the 5' sequences of were variable, but their 3' sequences were identical (Obata et al. 1998; Kohmura et al. 1998). After an extensive genomic analysis, they were found to be part of three clusters of nearly 60 genes in total. The clusters, which were named the α, β and γ protocadherins, are arrayed in tandem and span nearly 1MB (Sugino et al. 2000; Wu et al. 2001).

Through the work of several papers (Wu and Maniatis 2000; Tasic et al. 2002; Kehayova et al. 2011; Guo et al. 2012), a remarkable form of gene regulation was revealed for the clustered protocadherins. Within each cluster, there are numerous 5' "variable" exons and a few 3' "constant" exons. In all three clusters, each variable exon is preceded by an individual promoter. For the α and γ protocadherins, multiple large "variable" exons encode the complete sequence for six EC domains and part of a transmembrane domain. These exons then undergo cis-splicing with three "constant" exons that encodes for the remainder of the transmembrane domain and the cytoplasmic tail. For the βs, there are no "constant" exons. Instead, each protein is

entirely encoded by a single exon, and although their ICDs are not identical, they are highly conserved. As a result, each member within the α, β and or γ families have unique extracellular domains, but identical intracellular domains. Each variable exon has its own promoter 5' of the exon as well as a conserved sequence element (CSE) that is required for efficient transcription, and through a promoter choice mechanism, they are believe to be expressed in a stochastic manner (Tasic et al. 2002; Toyoda et al. 2014). Support for this has come from single-cell RT-PCR studies of Purkinje neurons, which showed that each cluster appears to be randomly and combinatorially expressed, and primarily in a monoallelic manner (Esumi et al. 2005; Kaneko et al. 2006; Hirano et al. 2012).

From these studies, it has been estimated that every neuron could express up to 10-15 different clustered protocadherins, generating high levels of surface diversity among neurons. However, this has never been properly or extensively characterized. Nevertheless, this arrangement gave much belief into a theoretical framework where the clustered protocadherins could provide sufficient neural surface diversity to wire the brain (Yagi 2012). This possibility drew immediate comparison to the Down Syndrome Cell Adhesion Molecule (Dscam) gene in *Drosophila*, which can generate 19,008 different extracellular domains via alternative splicing and functions as an important axon guidance molecule (Schmucker et al. 2000; Hummel et al. 2003). Dscam EC domains exhibit purely homophilic binding, meaning they will only bind to like-proteins (Wojtowicz et al. 2004; Wojtowicz et al. 2007), yet mediate repulsive behaviors between neurons (Wang et al. 2002). It is therefore believed that this is an elegant mechanism to

give neurons the ability to mediate self-recognition to avoid synapsing with itself and to efficiently sample other cells in the surrounding area. Since the DSCAM in mammals does not undergo alternative splicing like it does in *Drosophila* (Yamakawa et al. 1998; Agarwala et al. 2001), the clustered protocadherins were proposed to play the equivalent role of Dscams to provide an efficient mean to regulate self-avoidance.

7.3.2. Clustered Protocadherins: Evidence for Self-Avoidance

In the self-avoidance model, neurons or neurites that express the same complement of protocadherins and come into contact will transiently undergo binding mediated by the homophilic matching of the protocadherins. Since this is the exact opposite of self-avoidance, this binding event must then initiate intracellular signaling that inhibits the adhesive bonds, resulting in a repulsive signal. Strong *in vivo* evidence for this model comes from a study that conditionally knocked out all 22 gamma protocadherins in the retina and cerebellum, and examined the dendrites of starburst amacrine cells (SAC) and Purkinje cells (Lefebvre et al. 2012). The knockout mice displayed disrupted self-avoidance, causing the dendrites of the same neurons to cross over each other, whereas wild type dendrites never did. Further experiments revealed that this is not based on a particular combination of gamma isoforms, and that any one isoform is sufficient to produce self-avoidance (Lefebvre et al. 2012).

Studies on the olfactory system also lend support to the role in self-avoidance. Deletion of the constant region of alphas has been shown to disrupt axonal coalescence and targeting (Hasegawa et al. 2008), suggesting a role in self-recognition and targeting. A

recent paper deleted each cluster individually, and found minimal impact, whereas deletion of all 3 led to severe defects in dendritic arborization and loss of self-avoidance (Mountoufaris et al. 2017). Similar to the SAC study, self-avoidance was restored by over-expressing a single tri-cluster repertoire (i.e. 1 alpha, 1 beta, and 1 gamma), but targeting failed likely because each neuron recognized every neuron as self.

A similar but different method for maximizing neural wiring efficiency neural wiring is tiling, a process that maximizes target coverage of axons. Where self-avoidance is a mechanism to maximize the dendritic space from one soma by avoiding isoneuronal overlap, tiling is a mechanism by which neurons of the similar type or function avoid taking up overlapping space. In other words, self-avoidance repels sister neurites from the same cell, tiling repels neurites of the same cell type to regulate equal functional coverage. Protocadherin alpha-C2 has been found to be critical for this process in serotonergic neurons, and loss of it causes defects in tiling of serotonergic wiring and behavior defects (Chen et al. 2017).

Results of aggregation assays of clustered protocadherins in tissue culture have been interpreted to support the initial component of the self-avoidance model. In all cases tested, if two populations of cells that are transfected with different clustered protocadherins were not completely matched, coaggregation would not occur and they would segregate, forming completely separate aggregates (Schreiner and Weiner 2010; Thu et al. 2014). This demonstrates that combinations are important differentiating factors, and that the complement of isoforms expressed mediates homophilic adhesion. and several structural studies have defined the atomic characteristics of the adhesive

properties of the clustered protocadherins mediating this specificity (Rubinstein et al. 2015; Goodman et al. 2016a; Nicoludis et al. 2016; Goodman et al. 2016b). In the simplest scenario, sister neurites would express the same protocadherins, and cause them to bind homophilically. How then would they function to cause self-avoidance if matched repertoires induce adhesion? The prevailing model is that the adhesive interface is just the first step, and that a signaling event must cause them to “disengage” after the initial recognition event. Although it is known that the ICD of the protocadherins can be cleaved by gamma secretase (Buchanan et al. 2010), no model has been proposed for how the signaling to "disengage" after homophilic binding would occur.

7.3.3. Clustered Protocadherins: Evidence Against Self-Avoidance

Although the principles of self-avoidance and tiling have strong support in several cases, there are just as many examples that contradict these findings, or suggest different functions. The first knockout of the entire gamma cluster caused a decrease in spinal cord synapses and reduced mobility, however it also caused neonatal death due to neurodegeneration of spinal interneurons (Wang et al. 2002b). By crossing that mouse to a Bax knockout and also using a gamma hypomorph, the effect of apoptosis could be separated from the loss of the gammas (Weiner et al. 2005). The mice exhibited decreases in spinal cord synapses and profound neurological defects. No obvious defects in self-avoidance or tiling were observed, and so these studies suggest that instead, the gammas in the spinal cord play a role in development of synapse formation.

In the cortex, a conditional KO of all gammas did not result in cell death, nor reduce the density of their synapses. Instead, it caused a reduction in the complexity of dendrite arborizations (Garrett et al. 2012). A similar phenotype was observed for knockout and knockdowns of the alphas in the hippocampus and cultured hippocampal cells (Suo et al. 2012). These findings suggest that in the cortex and hippocampus, the gammas and alphas regulate the complexity of dendritic arborization instead of self-avoidance. In direct contrast to the self-avoidance model, it was found that dendritic complexity is regulated by homophilic matching of gammas with surrounding neurons and also astrocytes. Therefore, homophilic matching is actually required to promote dendritic complexity, and repulsive responses between cells with matching gammas was never observed in co-cultures (Molumby et al. 2016). It has been further shown that gammas regulate dendritic spine morphogenesis via blocking neurligin-neurexin1B interactions (Molumby et al. 2017), and not self-avoidance.

7.3.4. Clustered Protocadherins: Lessons of Context

Therefore, it can be seen that early generalizations that all protocadherins mediated self-avoidance were premature. The emerging theme is that different functions are observed depending on the location and context of the nervous system that is being studied, and this may be due to the different repertoire of isoforms they express. For example, early clustered protocadherin studies asserted that all neurons constitutively express all five C-type isoforms, and this was perpetuated for several years despite any evidence. Recent cell studies have now shown that this is not true for olfactory sensory neurons, nor in serotonergic neurons, and that serotonergic neurons primarily only

express one isoform, alpha-C2 (Chen et al. 2017; Mountoufaris et al. 2017). In other words, while some neurons appear to have tight regulation for which clustered protocadherins can be expressed, others have appear to have stochastic expression, albeit with slight elements of regulation.

Thus, the clustered protocadherins do not lend themselves to simple generalizations, and it is now abundantly clear that their functions are complex and highly dependent on the cellular and molecular context in which they are being studied and the specific complement being expressed. This idea extends to their tolerance for redundancy. For example in most neural systems, no one particular isoform is essential, and the required function relies on having a diverse pool of isoforms. In other cases, specific isoforms are necessary to mediate a cellular function (Chen et al. 2012; Chen et al. 2017). Many aspects of clustered protocadherin function remain to be understood. Ultimately, future work would build towards addressing a unified model of all their functions in neural wiring, and begin to understand how they interact with other protocadherins, and whether or not there is hierarchical ordering within their combinatorial functions and how that may change depending on which part of the nervous system is being examined.

7.3.5. δ-Protocadherins

Despite consisting of only nine members, the δ-protocadherins are less well understood than the clustered (α, β and γ) protocadherins. In contrast to the clustered protocadherins, they are not organized at a single genomic locus but instead are scattered throughout the genome. Currently, it is unknown how these genes are

regulated and more importantly, if they are co-regulated as a family or as independent genes. It is clear, however, that they play critical roles in numerous aspects of neural development. Although deltas are far outnumbered by the clustered protocadherins, the deltas show much less functional redundancy. Thus, whereas deleting one or even several clustered protocadherins can have no impact, removing or mutating a single δ-protocadherins can be significant. This suggests that on a gene per gene basis, the deltas may be more critical to neural development. While various aspects of the deltas in neural development have been revealed, much remains unknown about them. In particular, there is no clear model for how the family functions together.

7.3.6. δ-Protocadherins: Discovery and Genomic organization

The δ-protocadherins were discovered concomitantly with members of the clustered protocadherins (Sano et al. 1993; Yoshida et al. 1998; Strehl et al. 1998; Yoshida et al. 1999; Wolverton and Lalande 2001). After all the protocadherins were discovered, the δ-protocadherins were classified as a sub-family of the protocadherin family based on motifs (CM1 and CM2) in their intracellular domains (Vanhalst et al. 2005; Redies et al. 2005) that are not present in any other protocadherins. The δ-protocadherins are further divided into two sub-groups, the δ-1s (Pcdh 1, 7, 9 and 11X/Y), and the δ-2s (Pcdh 8, 10, 17, 18, and 19) based on the number of extracellular domains each gene contains. The delta-1 subfamily has seven extracellular cadherin-like domain repeats, while the delta-2 subfamily has only six, similar to the α, β and γ protocadherins. It is currently unknown what the functional consequences, if any, are for these differences in the number of extracellular domains. An additional distinction between the δ-1 and δ-2 is

the presence of the CM3 (RRVTF) motif in the intracellular domains of the δ-1 protocadherins. This motif is of particular interest because it allows the ICD to bind with protein phosphatase-1α (PP1α) (Yoshida et al. 1999; Zhou et al. 2017), which is involved in LTP (Munton et al. 2004) and therefore could contribute towards regulating synaptic plasticity.

Currently, nothing is known about the mechanisms by which δ-protocadherin expression is regulated. However, similar to the clustered protocadherins, the whole extracellular region of the δ-protocadherins is encoded by one large exon (roughly 2.5 kbs). This is in stark contrast to the classical cadherins, such as N-cadherin, which has a total of 16 exons, most of which are found in the region coding the extracellular domains (Miyatani et al. 1992). Interestingly, sequence alignments of the extracellular, intracellular, and full length proteins of the deltas reveals considerably low sequence homology for a protein family, and cursory alignments of sequences upstream of the transcription start site reveals no common motifs (personal observation).

7.3.7. δ-Protocadherins: Evolution

Understanding evolutionary relationships can provide deep insight into gene function. The expansion of genome sequencing to lower metazoan species allowed for large-scale genomic analysis of cadherin evolution to be performed. Although many aspects about protocadherin evolution are unknown, a few studies have provided information to establish a basic framework of their relationships. The overall relationship that has emerged is that the clustered protocadherins, which are only found in vertebrates, have

emerged much later in evolution than the δ-protocadherins (Hulpiau and Van Roy 2009; Hulpiau and Van Roy 2011; Hulpiau et al. 2013; Hulpiau et al. 2016), which have conserved sequences in deuterostomes, lophotrochozoa, and even the basal chordate *Cnidaria intestinalis* (Noda and Satoh 2008). A transcriptome analysis of extant jawless vertebrates (lamprey, cyclostomes) further supported this relationship by finding that cyclostomes lack the clustered protocadherins but have numerous δ-protocadherins (Ravi et al. 2016). Thus, it appears that the clustered protocadherins are evolutionary derived from the non-clustered protocadherins. This relationship is consistent with the lack of redundancy in δ-protocadherins, ands raises the possibility that origins in the clustered protocadherins may have resulted from gene duplications of one or more δ-protocadherins.

7.3.8. δ-Protocadherins: Expression and Putative Combinatorial Codes

Similar to clustered protocadherins, δ-protocadherin expression can be detected throughout the lifespan of the animal but are tightly regulated in a spatiotemporal manner. Usually, expression is most prominent during early embryonic stages, and then generally decreases to a steady level by adulthood (Hirano et al. 1999; Yamamoto et al. 2000; Gaitan and Bouchard 2006; Redies et al. 2008; Liu et al. 2009; Priddle and Crow 2013).

Two general themes emerge from the expression patterns of δ-protocadherins in the nervous system. First, they appear to be broadly expressed, yet differentially restricted to certain regions. For example, their expression in the mouse is under tight spatial

control for specific layers of the somatosensory cortex (Krishna-K et al. 2011), the visual cortex (Krishna-K et al. 2008), the spinal cord (Lin et al. 2012), the amygdala (Hertel et al. 2012), the basal ganglia (Hertel et al. 2008), and the cerebellum (Luckner et al. 2001). Many other regions and species have been examined (Hirano et al. 1999; Aoki et al. 2003; Müller et al. 2004; Vanhalst et al. 2005; Redies et al. 2005; Redies et al. 2008; Etzrodt et al. 2009; Liu et al. 2009; Blevins et al. 2011; Asahina et al. 2012), and essentially every region of the brain and nervous system exhibits spatiotemporal restriction of some deltas based on anatomical boundaries. This suggests that they play important roles in cell migration, and boundary formation of different regions.

Secondly, and consistent with the clustered protocadherins, is that they appear to be expressed in combination. This has been inferred from double label RNA *in situ* hybridization studies on various regions of the nervous system (Asahina et al. 2012; Krishna-K et al. 2011; 2008, Lin lab unpublished data). Since each delta overlaps with every other delta to a certain degree, it has been proposed that neurons expressed combinations of δ-protocadherins. However, this approach does not actually define what the combinations actually are, and a single-cell analysis of δ-protocadherins has never been reported. This leaves many important questions to be answered. For example, how many deltas can be expressed in a neuron? Are some more frequently expressed together than others, indicating the possibility of co-regulation? And most importantly, how does combinatorial expression affect their function? Since it is clear that neurons express at least two deltas, it is critical to understand how combinatorial

expression affects their adhesive properties to understand what role they play in neural circuit formation.

7.3.9. δ-Protocadherins Function as Adhesion Molecules

As members of the cadherin superfamily, δ-protocadherins are presumed to mediate calcium-dependent adhesion. To confirm this, immortalized cells are usually transfected with the cDNA of the gene of interest into immortalized cells, and then assayed for whether or not the expression induces cell aggregation when the cells are manipulated such that they are not allowed to adhere to the substratum. Several deltas have been confirmed to induce aggregation using mouse fibroblast L cells, which do not express endogenous cadherins. These include Pcdh1, Pcdh7, Pcdh9, Pcdh10 and Pcdh19 (Sano et al. 1993; Hirano et al. 1999; Yoshida 2003; Tai et al. 2010; Izuta et al. 2015). Pcdh19 has also been confirmed using Chinese Hamster Ovary cells (CHO), which are also void of endogenous cadherins, and beads coated with purified ectodomain protein (Emond et al. 2011). However, no reports have been made for Pcdh8, Pcdh9, Pcdh11x or Pcdh18 and no study has systematically assayed all the deltas within the same study.

Interestingly, after a study found that N-cadherin and Pcdh19 function together to regulate cell movements of anterior neurulation in zebrafish (Biswas et al. 2010) a follow up study found that Pcdh19 forms a complex with N-cadherin, and this complex creates a new adhesive specificity that no longer recognizes Pcdh19 or N-cadherin alone.

(Emond et al. 2011). This example emphasizes the need to understand how adhesive functions can change as a result of combinatorial expression.

7.3.10. δ-Protocadherins: Function in Early Stages of Neural Development

After the initiation of neural induction during gastrulation (Spemann and Mangold 1924), the embryo undergoes many changes as it begins to organize the nervous system. Several δ-protocadherins play important roles in these early stages of neural development. In the chicken, protocadherin 1 was found to be required for proper cell sorting and migration of neural crest cells to form dorsal root ganglia (Bononi et al. 2008). In *Xenopus*, the Protocadherin 1 paralog Axial protocadherin (AXPC) has been shown to mediate pre-notochord cell sorting (Kuroda et al. 2002), and also regulate cell fate during notochord morphogenesis (Yoder and Gumbiner 2011). Also in *Xenopus*, neural fold-protocadherin (NFPC, the paralog to Pcdh7) was shown to be required for ectodermal differentiation, and that expression of a dominant negative form disrupts the integrity of the ectoderm in the embryo (Bradley et al. 1998). In zebrafish, knockdown of pcdh19 prevents proper convergence of the anterior neural plate, resulting in disrupted morphogenesis of the brain (Emond et al. 2009), and pcdh9 establishes laminar embryo structures and early development of the retina (Izuta et al. 2015).

In the mouse, protocadherin 8 is required for proper somitogenesis (Rhee et al. 2003). The zebrafish and *Xenopus* homolog, Paraxial Protocadherin (PAPC) is required for proper convergence during gastrulation (Bouwmeester et al. 1996) in zebrafish and in *Xenopus*, it was first identified as a dorsal-specific marker in the embryo (Bouwmeester

et al. 1996). It was then found to mediate selective adhesion and cell movements of the mesoderm during gastrulation (Kim et al. 1998) and also cranial neural crest cell migration (Schneider et al. 2014). Many follow up studies detailed the mechanism, which involve several interactions with Rho GTPases, Sprouty, FLRT3, xANR5 and casein kinase 2B to modulate the cytoskeleton and the planar cell pathway, and down regulate C-cadherin (Medina et al. 2004; Chen and Gumbiner 2006; Chung et al. 2007; Wang et al. 2008; Chen et al. 2009; Kietzmann et al. 2012; Kai et al. 2015). In the mouse, Pcdh11x was shown in both *in vivo* and *in vitro* to regulate the differentiation and proliferation of neural stem cells (Zhang et al. 2014), and Pcdh9 has been identified as a downstream effector of Neurogenin2, a pro-neural transcription factor important in the developing neocortex (Mattar et al. 2004). These examples demonstrate the wide range of important functions deltas play in early neural development.

7.3.11. δ-Protocadherins: Functions in Axon Guidance, Tract Formation, and Synapse Regulation

δ-protocadherins also play critical roles at later stages of neural development. In cultured neurons of the mouse thalamus, Pcdh9 is detectable along the axon, growth cone, neurites, and synapse (Asahina et al. 2012). Pcdh10 and Pcdh8 have also been found to be localized to the synapse *in vivo* (Hirano et al. 1999; Yasuda et al. 2007). This suggests that they could be acting to regulate aspects of axon guidance, neurite extension, and synapse formation. Although the neuronal localization of other deltas remains to be characterized, it is expected that they would display similar patterns.

The limited studies done on delta functions support these possibilities. In the mouse, Pcdh10 null mice display defects in thalamacortical and corticothamalic axonal projections (Uemura et al. 2007). In mice over expressing Pcdh10, convergence of some olfactory sensory neuron populations is disrupted (Williams et al. 2011). In mouse neuronal cultures, Pcdh10 was found to regulate synapse elimination through the proteasome (Tsai et al. 2012). Pcdh17 has been found mediate group extension and fasciculation of axons from in the amygdala (Hayashi et al. 2014). In zebrafish, pcdh18b knockdowns displayed reduced axonal arborization in motor neurons (Biswas et al. 2014). In the zebrafish retina, NFPC regulates the initiation of axon elongation of retinal ganglia cells (Piper et al. 2008). Interestingly, although it was not found to function in intraretinal guidance itself, reduction of NFPC expression in RGCs prevented them from being sensitive to Netrin-1 (Leung et al. 2015), and is also upregulated at the growth cone by Sema3F to achieve accurate targeting (Leung et al. 2013). These studies emphasize the potential important in signaling cross talk between protocadherins and canonical axon guidance molecules.

The protocadherin 8 ortholog in rat was found to be activity regulated in the hippocampus by NMDA dependent synaptic LTP activity, and was therefore given the name Arcadlin (activity-regulated cadherin-like protein) (Yamagata et al. 1999). A follow up study found that it can regulate dendritic spine number by triggering N-cadherin endocytosis through a MAPK signaling pathway (Yasuda et al. 2007). Finally, in mouse cortical cultures, Protocadherin 11x expression has been observed to negatively regulate dendritic complexity though PI3K and AKT signaling pathways (Wu et al.

2015). Together, these studies support the role of δ-protocadherins in multiple aspects of neural wiring. However, a unified theme of their function remains to be developed.

7.3.12. δ-Protocadherins: Functions in Intracellular Signaling

Except for the conserved motifs, the deltas exhibit high variation in their intracellular cytoplasmic domains (ICD). This diversity is further enhanced by extensive alternative splicing of the cytoplasmic domains, producing several isoforms for each gene (Vanhalst et al. 2005; Blevins et al. 2011). The significance and regulation of isoform selection is not understood, but isoforms of Pcdh8 have been shown to be dynamic in the developing brain (Makarenkova et al. 2005), and isoforms of Pcdh11X/Y are localized to different sub-regions of the brain (Ahn et al. 2010). Together, these observations suggest a role for isoform-specific function. The intracellular signaling functions of the δ-protocadherins are largely unknown, but due to their variation are expected to have wide ranging capabilities compared to the classical cadherins and clustered protocadherins. The few studies that have identified ICD interactions support the wide range of possible functions. For example, the ICD of Pcdh7 was found to interact with protein phosphatase-1α through a yeast two-hybrid screen. This was confirmed via co-IP in mammalian cells, where Pcdh7 was also able to modulate the activity of PP1a (Yoshida et al. 1999). PP1a is known to be a regulator of long-term potentiation (LTP) (Mulkey et al. 1994; Malenka 1994), and therefore Pcdh7 may modulate synaptic plasticity through its ICD. In *Xenopus*, the ICD of PAPC interacts with Sprouty to inhibit convergent extension through tyrosine kinase signaling (Wang et al. 2008).

Since the protocadherins lack the β-catenin binding domain that allows for classical cadherins to bind to the cytoskeleton, they were initially thought to either not associate with the cytoskeleton, or do so through alternative mechanisms. An alternative mechanism began to be uncovered as Pcdh10, Pcdh17, Pcdh18b, and Pcdh19 were all independently found to associate with Nap1 (Nakao et al. 2008; Tai et al. 2010; Biswas et al. 2014; Hayashi et al. 2014). Nap1 is a component of the WAVE complex, which regulates actin dynamics and lamellipodia (Eden et al. 2002). A screen later identified many proteins that can link to the cytoskeleton via association with the WAVE complex and found that all δ-2s associate with the WAVE complex through a shared WIRS motif (Chen et al. 2014). What this means for functional differences between δ-1 and δ-2 remains to be seen.

7.3.13. δ-Protocadherins: Functions in Neurological Diseases

Genetic variations in δ-protocadherins have been implicated in several forms of neurological conditions. Through numerous genome wide association studies (GWAS), several deltas have been associated with autism, including Pcdh8 (Butler et al. 2015), Pcdh9 (Bucan et al. 2009; Hussman et al. 2011; Luo et al. 2012; Girirajan et al. 2013) and Pcdh10 (Morrow et al. 2008; Tsai et al. 2012). A mouse model null for Pcdh9 displayed autism like behaviors (Bruining et al. 2015), and a mouse model of Pcdh10 found that heterozygous males have altered social behavior, and females have alterations in motor coordination. Most interestingly, dendritic spine density was found to be increased in parts of the amygdala (Schoch et al. 2016). This is consistent with the

role of Pcdh10 in synapse elimination (Tsai et al. 2012), and suggests one possible mechanism underlying autism is altered regulation of synaptic densities.

In recent years, protocadherin 19 has received the most attention for its causative role in Protocadherin 19 Female Limited Epilepsy (PCDH19 FLE), and is now considered a major epilepsy gene. Although first observed in 1971 (Juberg and Hellman 1971), it wasn't until 2008 that protocadherin 19 was identified as the underlying causative gene (Dibbens et al. 2008). Unlike channelopathies which have obvious functional implications to epilepsy, cell-adhesion molecules such as Pcdh19 do not, and the mechanism in FLE remains to be determined. Understanding PCDH19 FLE is further complicated by its unique mode of inheritance. Since Pcdh19 is located on the X chromosome, it follows an X-linked mode of transmission, yet while males can be carriers of mutant alleles, they appear to never be affected unless mosaic. This has caused a "cellular interference" model to be proposed. In this model, neurons in males would be all mutant or all wild type based on their single inherited allele. In contrast, females, who would inherit two alleles, would have neurons randomly expressing wild-type and mutant alleles due to X-inactivation (Depienne et al. 2009; Hynes et al. 2010; Depienne et al. 2010). This mixed population may cause alterations to neural wiring, and lead to the features of PCDH19 FLE. A mouse model with exons 1-3 of Pcdh19 removed showed no apparent signs of impaired brain development, nor epilepsy (Pederick et al. 2016). Although FLE is being increasingly characterized in the clinical setting (Specchio et al. 2011; Higurashi et al. 2012; Marini et al. 2012; Higurashi et al. 2013; Leonardi et al. 2014; Cappelletti et al. 2015; Gagliardi et al. 2015; Trivisano et al.

2016), there is no current treatment. This is largely because the function of Pcdh19 is not well understood, and therefore how mutations could lead to these symptoms is unclear. However, recent work has found potential links to steroid and hormonal imbalances (Tan et al. 2015; Trivisano et al. 2017), suggesting a possible avenue for treatment. In addition to PCDH19 FLE, Pcdh7 was recently determined to be highly associated with common epilepsies in a large GWAS (Consortium 2014).

Deltas are also implicated with conditions that are rarer. For example, Pcdh8 is associated with psychomotor delay (Castéra et al. 2013) and schizophrenia (Bray et al. 2002). Genetic variants in Pcdh11X were found to be associated with late onset Alzheimer's (Carrasquillo et al. 2009), although three follow up studies on different populations failed to detect the same association (Lescai et al. 2010; Beecham et al. 2010; Miar et al. 2011) suggesting that the initial result was population specific, or the effect was over-estimated. Pcdh11 is also implicated in rarer conditions such as Klinefelter's Syndrome (XXY) (Ross et al. 2006), dyslexia (Veerappa et al. 2013), and non-syndromic language delay (Speevak and Farrell 2011).

7.3.14. δ-Protocadherins: Functions in Non-Neurological Diseases and Conditions

δ-Protocadherins are also implicated in diseases and conditions outside the nervous system, showing their importance beyond neural development. These conditions are diverse, ranging from asthma (Toncheva et al. 2012) and bronchial hyperresponsiveness (Koppelman et al. 2009) to eczema (Koning et al. 2012). Most commonly, however, the deltas are associated with various forms of cancer. In many

cases, tumors have lower mRNA or protein expression levels and/or exhibit hypermethylation of the delta promoter (Beukers et al. 2013; Bujko et al. 2015; Lin et al. 2016). Usually, these decreases are associated with higher tumor grade, larger tumor size, and shorter recurrence free survival, suggesting that their expression status could be used as a biomarker. This type of association is the most common, and suggests that the deltas may play a tumor suppressor role. However, in some tumors, the exact opposite is found, and delta expression is enhanced (Lin et al. 2013; Zhou et al. 2017) and in one case, this increase was found to potentiate important cancer pathways such as ERK signaling (Zhou et al. 2017). These inconsistencies highlight the complex nature in which protocadherin function must be interpreted, and are consistent with the principle that different functions of deltas may stem from different cellular contexts. This pattern of cancer involvement again raises another interesting question. Why are many δ-protocadherins associated with cancer (Pcdh 1, 7, 8, 9, 10, 11X, and 17), while others are not (Pcdh 11Y, 18 and 19)? Also, since multiple deltas appear to be expressed in cells, why do they associate with cancer as only singular genes? Uncovering their expression patterns and regulation in the nervous system may shed more light into their combinatorial function in cancer, and vice versa.

7.3.15. δ-Protocadherins: Remaining Questions

It can be seen that the overall nature of the δ-protocadherins is highly diverse. Although there are some common themes among them, no central principle explaining δ-protocadherin function has emerged. Like the clustered protocadherins, it appears that the context in which they are studied in is important, and this aspect contributes to their

varied functional capabilities. This is likely due to the combinatorial nature of protocadherins, as well as interactions with other proteins such as classical cadherins and various canonical axon guidance molecules.

Many important questions need to be answered to better understand delta function in the nervous system. First, is there really a combinatorial code by which adhesive specificities can be defined? If so, to what extent are they expressed in combinations, and are some deltas co-expressed more than others? Insight into this could help understand why some deltas appear to be more important that others in particular neural systems if other deltas are also co-expressed in the same cells.

If they are expressed in combinations, a critical question to answer is then what does combinatorial expression do to their adhesive specificities or functions? It is important to determine what the purpose of combinations are. If they serve to act as agents of neural diversity, where the specific repertoire does not seem to matter, then they would function similarly to the clustered protocadherins. However, if the makeup with the combinatorial repertoire matters, then that would present a new mechanism by which combinatorial adhesive qualities are modulated. Addressing this issue will provide a better context to understand δ-protocadherin function.

A major limitation in previous studies in preventing a unified principle of delta function from being depicted is that they have never systematically been studied together before. This thesis aims to systematically address these questions by first developing a single-cell method to define combinatorial expression in single neurons, and then study the

adhesive properties of δ-protocadherin proteins in combinations. Together, they will provide a framework for which δ-protocadherin function can be further studied and understood in combinatorial contexts.

8. A New Approach Towards Understanding the Combinatorial Expression and Function of Neural Wiring Genes

In efforts to work towards building a comprehensive understanding of brain development, genomic and genetic approaches ultimately suffer from a major limitation - neither approach can offer a complete functional explanation of how the complexities of neural wiring occur through the concerted work of many genes. On one hand, genomic techniques such as microarrays or RNA-seq can describe or define expression patterns of specific cells or aspects of developmental dynamics by examining the whole transcriptome. However, in itself this approach offers little to no insight into what the *functional* mechanistic consequences are of the thousands of genes being expressed together to influence neural wiring. On the other hand, genetic techniques can reveal detailed information about the function of one, or perhaps two genes, but can never take into account the functional interactions of the many other genes also being expressed and influencing cellular behavior. As a consequence, our current knowledge of axon guidance and synaptogenesis is limited by views these two polar extremes provide. Our understanding of neural wiring will never improve unless new approaches are used to overcome this barrier.

How could this be achieved? The gap that occupies the space between genetic and genomic studies must begin to be filled in with knowledge that is able to explain function. Ultimately, the goal would to be able to fully integrate genomic and genetic data into a meaningful continuum that explains the function of genes involved in neural wiring when expressed in combinations. This will obviously not be achieved quickly nor easily, and will require being completed in a modular fashion. For example, to start, an ideal case would be taking a small gene family that exhibits some degree of combinatorial expression in neurons. If the genes are related enough in function such that they can be studied with the same assay, one could then test how different combinations affect their functional output. This could be considered one "module" consisting of a gene family or a group of similar genes. Already, this will provide more information than genetic studies alone because it encompasses multiple genes and reveals how their combinatorial expression influences their function. However, it is still extremely far away from approaching the scale of the genome, or even the ~200 genes involved in neural wiring. To extend this module to cover more genes, the same approach could then be performed with another group of genes gene family to create a second "module". These two modules could then be combined through more experiments to understand how different gene families influence function when expressed in combinations. This process could be performed over many groups of genes and gene families to produce a connected understanding of combinatorial gene function and possibly reveal the mechanistic principles behind Cajal's "chemotropism"

and Sperry's "chemoaffinity hypothesis" and represent a new paradigm for understanding neural development.

In summary, genomic studies do not provide enough detail on gene function, genetic studies do not study enough genes to reveal the complexities found *in vivo*, and currently there is no "middle ground" to begin connecting the two. The aim of this thesis is to demonstrate the utility of integrating expression and function studies of multiple genes in such a manner to produce a more meaningful interpretation into how combinatorial gene expression relates to the functional consequence of its gene products. To do so, I will study a small gene family of cell-adhesion molecules, the δ-protocadherins, which have been implicated in various aspects of neural development and wiring but how they function together is unknown. Studying the δ-protocadherins with this method will demonstrate how this new approach can be effective in increasing our knowledge of neural wiring.

9. References

Aboitiz F, Scheibel AB, Fisher RS, Zaidel E (1992) Fiber composition of the human corpus callosum. Brain research. doi: 10.1016/0006-8993(92)90178-C

Agarwala KL, Ganesh S, Amano K, et al (2001) DSCAM, a Highly Conserved Gene in Mammals, Expressed in Differentiating Mouse Brain. Biochemical and Biophysical Research Communications 281:697–705. doi: 10.1006/bbrc.2001.4420

Ahn K, Huh J-W, Kim D-S, et al (2010) Quantitative analysis of alternative transcripts of human PCDH11X/Y genes. American journal of medical genetics Part B, Neuropsychiatric genetics: the official publication of the International Society of Psychiatric Genetics 153B:736–744. doi: 10.1002/ajmg.b.31041

Aoki E, Kimura R, Suzuki ST, Hirano S (2003) Distribution of OL-protocadherin protein in correlation with specific neural compartments and local circuits in the postnatal mouse brain. Neuroscience 117:593–614.

Aoto J, Ting P, Maghsoodi B, et al (2007) Postsynaptic ephrinB3 promotes shaft glutamatergic synapse formation. The Journal of neuroscience : the official journal of the Society for Neuroscience 27:7508–7519. doi: 10.1523/JNEUROSCI.0705-07.2007

Asahina H, Masuba A, Hirano S, Yuri K (2012) Distribution of protocadherin 9 protein in the developing mouse nervous system. Journal of Neuroscience 225:88–104. doi: 10.1016/j.neuroscience.2012.09.006

Azevedo FAC, Carvalho LRB, Grinberg LT, et al (2009) Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. The Journal of comparative neurology 513:532–541. doi: 10.1002/cne.21974

Bakken TE, Miller JA, Luo R, et al (2015) Spatiotemporal dynamics of the postnatal developing primate brain transcriptome. Human Molecular Genetics 24:4327–4339. doi: 10.1093/hmg/ddv166

Bate CM (1976) Pioneer neurones in an insect embryo. Nature 260:54–56. doi: 10.1038/260054a0

Beecham, G.W., Naj, A.C., Gilbert, J.R., Haines, J.L., Buxbaum, J.D., and Pericak-Vance, M.A. (2010). PCDH11X variation is not associated with late-onset Alzheimer disease susceptibility. Psychiatric Genetics 20, 321–324.

Ben Chih, Engelman H, Scheiffele P (2005) Control of Excitatory and Inhibitory Synapse Formation by Neuroligins. Science 307:1324–1328. doi: 10.1126/science.1107470

Bentley D, Caudy M (1983) Pioneer axons lose directed growth after selective killing of guidepost cells. Nature 304:62–65.

Beukers W, Hercegovac A, Vermeij M, et al (2013) Hypermethylation of the Polycomb Group Target Gene PCDH7 in Bladder Tumors from Patients of All Ages. JURO 190:311–316. doi: 10.1016/j.juro.2013.01.078

Biederer T, Sara Y, Mozhayeva M, et al (2002) SynCAM, a synaptic adhesion molecule that drives synapse assembly. Science 297:1525–1531. doi: 10.1126/science.1072356

Biswas S, Emond MR, Duy PQ, et al (2014) Protocadherin-18b interacts with Nap1 to control motor axon growth and arborization in zebrafish. Molecular Biology of the Cell 25:633–642. doi: 10.1091/mbc.E13-08-0475

Biswas S, Emond MR, Jontes JD (2010) Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. The Journal of cell biology 191:1029–1041. doi: 10.1083/jcb.201007008

Bixby JL, Zhang R (1990) Purified N-cadherin is a potent substrate for the rapid induction of neurite outgrowth. The Journal of cell biology 110:1253–1260.

Blevins CJ, Emond MR, Biswas S, Jontes JD (2011) Differential expression, alternative splicing, and adhesive properties of the zebrafish δ1-protocadherins. Neuroscience 199:523–534. doi: 10.1016/j.neuroscience.2011.09.061

Bonkowsky JL, Yoshikawa S, O'Keefe DD, et al (1999) Axon routing across the midline controlled by the Drosophila Derailed receptor. Nature 402:540–544. doi: 10.1038/990122

Bononi J, Cole A, Tewson P, et al (2008) Chicken protocadherin-1 functions to localize neural crest cells to the dorsal root ganglia during PNS formation. Mechanisms of Development 125:1033–1047. doi: 10.1016/j.mod.2008.07.007

Bourgeois JP, Goldman-Rakic PS, Rakic P (1994) Synaptogenesis in the prefrontal cortex of rhesus monkeys. Cerebral Cortex 4:78–96.

Bouwmeester T, Kim S, Sasai Y, et al (1996) Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. Nature 382:595–601. doi: 10.1038/382595a0

Brackenbury R, Thiery JP, Rutishauser U, Edelman GM (1977) Adhesion among neural cells of the chick embryo. I. An immunological assay for molecules involved in cell-cell binding. Journal of Biological Chemistry 252:6835–6840.

Bradley RS, Espeseth A, Kintner C (1998) NF-protocadherin, a novel member of the cadherin superfamily, is required for Xenopus ectodermal differentiation. Current biology 8:325–334. doi: 10.1016/s0960-9822(98)70132-0

Bray NJ, Kirov G, Owen RJ, et al (2002) Screening the human protocadherin 8 (PCDH8) gene in schizophrenia. Genes, Brain and Behavior 1:187–191. doi: 10.1034/j.1601-183X.2002.10307.x

Brose K, Bland KS, Wang KH, et al (1999) Slit Proteins Bind Robo Receptors and Have an Evolutionarily Conserved Role in Repulsive Axon Guidance. Cell 96:795–806. doi: 10.1016/S0092-8674(00)80590-5

Bruining H, Matsui A, Oguro-Ando A, et al (2015) Genetic Mapping in Mice Reveals the Involvement of Pcdh9 in Long-Term Social and Object Recognition and Sensorimotor Development. Biological Psychiatry 78:485–495. doi: 10.1016/j.biopsych.2015.01.017

Bruner J, Tauc L (1966) Habituation at the synaptic level in Aplysia. Nature 210:37–39.

Bucan M, Abrahams BS, Wang K, et al (2009) Genome-Wide Analyses of Exonic Copy Number Variants in a Family-Based Study Point to Novel Autism Susceptibility Genes. PLOS Genet 5:e1000536. doi: 10.1371/journal.pgen.1000536

Bujko M, Kober P, Mikula M, et al (2015) Expression changes of cell-cell adhesion-related genes in colorectal tumors. Oncology letters 9:2463–2470. doi: 10.3892/ol.2015.3107

Burkhardt C, Müller M, Badde A, Garner CC, Gundelfinger ED, and Püschel AW (2005). Semaphorin 4B interacts with the post-synaptic density protein PSD-95/SAP90 and is recruited to synapses through a C-terminal PDZ-binding motif. FEBS Lett. *579*, 3821–3828.

Butler M, Rafi S, Hossain W, et al (2015) Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. International journal of molecular sciences 16:1312–1335. doi: 10.3390/ijms16011312

Cajal SR y (1909) Histologie du Système nerveux de l'Homme et des Vertebres, Trad. doi: 10.5962/bhl.title.48637

Campbell DS, Stringham SA, Timm A, et al (2007) Slit1a inhibits retinal ganglion cell arborization and synaptogenesis via Robo2-dependent and -independent pathways. Neuron 55:231–245. doi: 10.1016/j.neuron.2007.06.034

Cappelletti S, Specchio N, Moavero R, et al (2015) Cognitive development in females with PCDH19 gene-related epilepsy. Epilepsy & Behavior 42:36–40. doi: 10.1016/j.yebeh.2014.10.019

Carrasquillo MM, Zou F, Pankratz VS, et al (2009) Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. Nature genetics 41:192–198. doi: 10.1038/ng.305

Castellucci V, Pinsker H, Kupfermann I, Kandel ER (1970) Neuronal Mechanisms of Habituation and Dishabituation of the Gill-Withdrawal Reflex in Aplysia. Science 167:1745–1748. doi: 10.1126/science.167.3926.1745

Castéra L, Dehainault C, Michaux D, et al (2013) Fine mapping of whole RB1 gene deletions in retinoblastoma patients confirms PCDH8 as a candidate gene for psychomotor delay. European journal of human genetics : EJHG 21:460–464. doi: 10.1038/ejhg.2012.186

Caton A, Hacker A, Naeem A, et al (2000) The branchial arches and HGF are growth-promoting and chemoattractant for cranial motor axons. Development 127:1751–1766.

Chalasani SH, Sabelko KA, Sunshine MJ, et al (2003) A chemokine, SDF-1, reduces the effectiveness of multiple axonal repellents and is required for normal axon pathfinding. The Journal of neuroscience : the official journal of the Society for Neuroscience 23:1360–1371.

Challacombe JF, Snow DM, Letourneau PC (1996) Actin filament bundles are required for microtubule reorientation during growth cone turning to avoid an inhibitory guidance cue. Journal of Cell Science 109:2031–2040.

Chen B, Brinkmann K, Chen Z, et al (2014) The WAVE Regulatory Complex Links Diverse Receptors to the Actin Cytoskeleton. Cell 156:195–207. doi: 10.1016/j.cell.2013.11.048

Chen WV, Alvarez FJ, Lefebvre JL, et al (2012) Functional Significance of Isoform Diversification in the Protocadherin Gamma Gene Cluster. Neuron 75:402–409. doi: 10.1016/j.neuron.2012.06.039

Chen WV, Nwakeze CL, Denny CA, et al (2017) Pcdhαc2 is required for axonal tiling and assembly of serotonergic circuitries in mice. Science 356:406–411. doi: 10.1126/science.aal3231

Chen X, Gumbiner BM (2006) Paraxial protocadherin mediates cell sorting and tissue morphogenesis by regulating C-cadherin adhesion activity. The Journal of cell biology 174:301–313. doi: 10.1083/jcb.200602062

Chen X, Koh E, Yoder M, Gumbiner BM (2009) A protocadherin-cadherin-FLRT3 complex controls cell adhesion and morphogenesis. PLOS ONE 4:e8411. doi: 10.1371/journal.pone.0008411

Chisholm A, Tessier-Lavigne M (1999) Conservation and divergence of axon guidance mechanisms. Current opinion in neurobiology 9:603–615. doi: 10.1016/S0959-4388(99)00021-5

Chung HA, Yamamoto TS, Ueno N (2007) ANR5, an FGF target gene product, regulates gastrulation in Xenopus. CURBIO 17:932–939. doi: 10.1016/j.cub.2007.04.034

Colantuoni C, Lipska BK, Ye T, et al (2011) Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 478:519–523. doi: 10.1038/nature10524

Colón-Ramos DA, Margeta MA, Shen K (2007) Glia promote local synaptogenesis through UNC-6 (netrin) signaling in *C. elegans*. Science 318:103–106. doi: 10.1126/science.1143762

Consortium T (2014) Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. The Lancet Neurology. doi: 10.1016/S1474-4422(14)70171-1

Contractor A, Rogers C, Maron C, et al (2002) Trans-Synaptic Eph Receptor-Ephrin Signaling in Hippocampal Mossy Fiber LTP. Science 296:1864–1869. doi: 10.1126/science.1069081

Corriveau RA, Shatz CJ, Nedivi E (1999) Dynamic Regulation of cpg15 during Activity-Dependent Synaptic Development in the Mammalian Visual System. Journal of Neuroscience 19:7999–8008.

Curtis A (1961) Timing mechanisms in the specific adhesion of cells. Experimental cell research 8:107–122. doi: 10.1016/0014-4827(61)90343-3

Dalva MB, Takasu MA, Lin MZ, et al (2000) EphB Receptors Interact with NMDA Receptors and Regulate Excitatory Synapse Formation. Cell 103:945–956. doi: 10.1016/S0092-8674(00)00197-5

Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, Hayden Gephart MG, Barres BA, and Quake SR (2015). A survey of human brain transcriptome diversity at the single cell level. Proc. Natl. Acad. Sci. U.S.A. *112*, 7285–7290.

Deakin JE, Graves JAM, Rens W (2012) The evolution of marsupial and monotreme chromosomes. Cytogenetic and Genome Research 137:113–129. doi: 10.1159/000339433

Depienne C, Bouteiller D, Keren B, et al (2009) Sporadic Infantile Epileptic Encephalopathy Caused by Mutations in PCDH19 Resembles Dravet Syndrome but Mainly Affects Females. PLOS Genet 5:NaN–NaN. doi: 10.1371/journal.pgen.1000381

Depienne C, Trouillard O, Bouteiller D, et al (2010) Mutations and deletions in PCDH19 account for various familial or isolated epilepsies in females. Human Mutation 32:E1959. doi: 10.1002/humu.21373

Dibbens LM, Tarpey PS, Hynes K, et al (2008) X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. Nature genetics 40:776–781. doi: 10.1038/ng.149

Dickson BJ (2002) Molecular mechanisms of axon guidance. Science 298:1959–1964. doi: 10.1126/science.1072165

Dominici C, Moreno-Bravo JA, Puiggros SR, et al (2017) Floor-plate-derived netrin-1 is dispensable for commissural axon guidance. Nature 545:350–354. doi: 10.1038/nature22331

Duan X, Krishnaswamy A, la Huerta I De, Sanes JR (2014) Type II cadherins guide assembly of a direction-selective retinal circuit. Cell 158:793–807. doi: 10.1016/j.cell.2014.06.047

Duguay D, Foty RA, Foty R, et al (2003) Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. 253:309–323.

Easter SS, Ross LS, Frankfurter A (1993) Initial tract formation in the mouse brain. The Journal of neuroscience : the official journal of the Society for Neuroscience 13:285–299.

Eden S, Rohatgi R, Podtelejnikov AV, et al (2002) Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. Nature 418:790–793. doi: 10.1038/nature00859

Einheber, S., Schnapp, L.M., and Salzer, J.L. (1996). Regional and ultrastructural distribution of the α8 integrin subunit in developing and adult rat brain suggests a role in synaptic function. Journal of Comparative Neurology 105–134.

Emond, M.R., Biswas, S., Blevins, C.J., and Jontes, J.D. (2011). A complex of Protocadherin-19 and N-cadherin mediates a novel mechanism of cell adhesion. J. Cell Biol. 195, 1115–1121.

Emond MR, Biswas S, and Jontes JD (2009). Protocadherin-19 is essential for early steps in brain morphogenesis. Developmental Biology 334, 72–83.

Engle EC (2010) Human genetic disorders of axon guidance. Cold Spring Harbor perspectives in biology 2:a001784. doi: 10.1101/cshperspect.a001784

Essmann CL, Martinez E, Geiger JC, et al (2008) Serine phosphorylation of ephrinB2 regulates trafficking of synaptic AMPA receptors. Nature neuroscience 11:1035–1043. doi: 10.1038/nn.2171

Esumi S, Kakazu N, Taguchi Y, et al (2005) Monoallelic yet combinatorial expression of variable exons of the protocadherin-alpha gene cluster in single neurons. Nature genetics 37:171–176. doi: 10.1038/ng1500

Etzrodt J, Krishna-K K, Redies C (2009) Expression of classic cadherins and delta-protocadherins in the developing ferret retina. BMC Neuroscience 10:153. doi: 10.1186/1471-2202-10-153

Faissner A, Kruse J (1990) J1/tenascin is a repulsive substrate for central nervous system neurons. Neuron 5:627–637.

Fan J, Raper JA (1995) Localized collapsing cues can steer growth cones without inducing their full collapse. Neuron 14:263–274.

Flower WH (1865) On the Commissures of the Cerebral Hemispheres of the Marsupialia and Mono-Tremata as Compared with Those of the Placental Mammals. Philosophical Transactions of the Royal Society of London 155:633–651. doi: 10.2307/108897

Foty, R.A., and Steinberg, M.S. (2005). The differential adhesion hypothesis: a direct evaluation. Developmental Biology *278*, 255–263. doi: 10.1016/j.ydbio.2004.11.012

Fowler DK, Peters JH, Williams C, Washbourne P (2017) Redundant Postsynaptic Functions of SynCAMs 1–3 during Synapse Formation. Frontiers in molecular neuroscience 10:1227. doi: 10.3389/fnmol.2017.00024

Friedman HV, Bresler T, Garner CC, Ziv NE (2000) Assembly of new individual excitatory synapses: time course and temporal order of synaptic molecule recruitment. Neuron 27:57–69. doi: 10.1016/j.celrep.2012.12.005

Fukata Y, Fukata M (2017) Epilepsy and synaptic proteins. Current opinion in neurobiology 45:1–8. doi: 10.1016/j.conb.2017.02.001

Gagliardi M, Annesi G, Sesta M, et al (2015) PCDH19 mutations in female patients from Southern Italy. Seizure: European Journal of Epilepsy 24:118–120. doi: 10.1016/j.seizure.2014.08.010

Gaitan Y, Bouchard M (2006) Expression of the delta-protocadherin gene Pcdh19 in the developing mouse embryo. Gene expression patterns : GEP 6:893–899. doi: 10.1016/j.modgep.2006.03.001

Galtsoff PS (1923) The Amœboid Movement of Dissociated Sponge Cells. Biological Bulletin 45:153–161. doi: 10.2307/1536625

García-Alonso L, Fetter RD, Goodman CS (1996) Genetic analysis of Laminin A in Drosophila: extracellular matrix containing laminin A is required for ocellar axon pathfinding. Development 122:2611–2621.

Garrett AM, Schreiner D, Lobas MA, Weiner JA (2012) γ-protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. Neuron 74:269–276. doi: 10.1016/j.neuron.2012.01.028

Gazzaniga, M.S. (2000). Cerebral specialization and interhemispheric communication: does the corpus callosum enable the human condition? Brain *123*, 1293–1326.

Genç B, Özdinler PH, Mendoza AE, Erzurumlu RS (2004) A Chemoattractant Role for NT-3 in Proprioceptive Axon Guidance. PLoS biology 2:e403. doi: 10.1371/journal.pbio.0020403

Girirajan S, Johnson RL, Tassone F, et al (2013) Global increases in both common and rare copy number load associated with autism. Human Molecular Genetics 22:2870–2880. doi: 10.1093/hmg/ddt136

Giuditta A, Dettbarn WD, Brzin M (1968) Protein Synthesis in the Isolated Giant Axon of the Squid. Proceedings of the National Academy of Sciences of the United States of America 59:1284–1287. doi: 10.2307/58405

Godt D, Tepass U (1998) Drosophila oocyte localization is mediated by differential cadherin-based adhesion. Nature 395:387–391. doi: 10.1038/26493

Goodman KM, Rubinstein R, Thu CA, et al (2016a) Structural Basis of Diverse Homophilic Recognition by Clustered α- and β-Protocadherins. Neuron 90:709–723. doi: 10.1016/j.neuron.2016.04.004

Goodman KM, Rubinstein R, Thu CA, et al (2016b) γ-Protocadherin structural diversity and functional implications. eLife 5:e20930. doi: 10.7554/eLife.20930

Gordon L, Mansh M, Kinsman H, Morris AR (2010) Xenopus sonic hedgehog guides retinal axons along the optic tract. Developmental Dynamics 239:2921–2932. doi: 10.1002/dvdy.22430

Graves J (1996) Mammals that break the rules: genetics of marsupials and monotremes. Annual review of genetics 30:233–260. doi: 10.1146/annurev.genet.30.1.233

Greenough WT, Klintsova AY, Irwin SA, et al (2001) Synaptic regulation of protein synthesis and the fragile X protein. Proceedings of the National Academy of Sciences 98:7101–7106. doi: 10.1073/pnas.141145998

Grenningloh G, Rehm EJ, Goodman CS (1991) Genetic analysis of growth cone guidance in Drosophila: fasciclin II functions as a neuronal recognition molecule. Cell 67:45–57.

Grunwald IC, Korte M, Wolfer D, et al (2001) Kinase-independent requirement of EphB2 receptors in hippocampal synaptic plasticity. Neuron 32:1027–1040.

Grutzendler J, Kasthuri N, Gan W-B (2002) Long-term dendritic spine stability in the adult cortex. Nature 420:812–816. doi: 10.1038/nature01276

Gundersen RW, Barrett JN (1979) Neuronal chemotaxis: chick dorsal-root axons turn toward high concentrations of nerve growth factor. Science 206:1079–1080. doi: 10.1126/science.493992

Gundersen RW, Barrett JN (1980) Characterization of the turning response of dorsal root neurites toward nerve growth factor. The Journal of cell biology 87:546–554.

Guo H-F, Vander Kooi CW (2015) Neuropilin Functions as an Essential Cell Surface Receptor. The Journal of biological chemistry 290:29120–29126. doi: 10.1074/jbc.R115.687327

Guo Y, Monahan K, Wu H, et al (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. Proceedings of the National Academy of Sciences of the United States of America 109:21081–21086. doi: 10.1073/pnas.1219280110

Harris AK (1976) Is Cell sorting caused by differences in the work of intercellular adhesion? A critique of the Steinberg hypothesis. Journal of theoretical biology 61:267–285. doi: 10.1016/0022-5193(76)90019-9

Harrison RG (1910) The outgrowth of the nerve fiber as a mode of protoplasmic movement. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology 9:787–846. doi: 10.1002/jez.1400090405

Hasegawa S, Hamada S, Kumode Y, et al (2008) The protocadherin-α family is involved in axonal coalescence of olfactory sensory neurons into glomeruli of the olfactory bulb in mouse. Molecular and Cellular Neuroscience 38:66–79. doi: 10.1016/j.mcn.2008.01.016

Hatta K, Takeichi M (1986) Expression of N-cadherin adhesion molecules associated with early morphogenetic events in chick development. 320:447–449. doi: 10.1038/320447a0

Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 489:391–399. doi: 10.1038/nature11405

Hayano Y, Sasaki K, Ohmura N, et al (2014) Netrin-4 regulates thalamocortical axon branching in an activity-dependent fashion. Proceedings of the National Academy of Sciences of the United States of America 111:15226–15231. doi: 10.1073/pnas.1402095111

Hayashi S, Inoue Y, Kiyonari H, et al (2014) Protocadherin-17 mediates collective axon extension by recruiting actin regulator complexes to interaxonal contacts. Developmental cell 30:673–687. doi: 10.1016/j.devcel.2014.07.015

Hedgecock EM, Culotti JG, Hall DH (1990) The unc-5, unc-6, and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in *C. elegans*. Neuron 4:61–85. doi: 10.1016/0896-6273(90)90444-K

Henkemeyer M, Itkis OS, Ngo M, et al (2003) Multiple EphB receptor tyrosine kinases shape dendritic spines in the hippocampus. The Journal of cell biology 163:1313–1326. doi: 10.1083/jcb.200306033

Herculano-Houzel S, Mota B, Lent R (2006) Cellular scaling rules for rodent brains. Proceedings of the National Academy of Sciences of the United States of America 103:12138–12143. doi: 10.1073/pnas.0604911103

Hermey G, Mahlke C, Gutzmann JJ, et al (2013) Genome-Wide Profiling of the Activity-Dependent Hippocampal Transcriptome. PLOS ONE 8:e76903. doi: 10.1371/journal.pone.0076903

Hertel N, Krishna-K, Nuernberger M, Redies C (2008) A cadherin-based code for the divisions of the mouse basal ganglia. The Journal of comparative neurology 508:511–528. doi: 10.1002/cne.21696

Hertel N, Redies C, Medina L (2012) Cadherin expression delineates the divisions of the postnatal and adult mouse amygdala. The Journal of comparative neurology 520:3982–4012. doi: 10.1002/cne.23140

Higurashi, N., Nakamura, M., Sugai, M., Ohfu, M., Sakauchi, M., Sugawara, Y., Nakamura, K., Kato, M., Usui, D., Mogami, Y., et al. (2013). PCDH19-related female-limited epilepsy: further details regarding early clinical features and therapeutic efficacy. Epilepsy Res. 106, 191–199.

Higurashi N, Shi X, Yasumoto S, et al (2012) PCDH19 mutation in Japanese females with epilepsy. Epilepsy research 99:28–37. doi: 10.1016/j.eplepsyres.2011.10.014

Hirano K, Kaneko R, Izawa T, et al (2012) Single-neuron diversity generated by Protocadherin-β cluster in mouse central and peripheral nervous systems. 5:90. doi: 10.3389/fnmol.2012.00090

Hirano S, Yan Q, Suzuki ST (1999) Expression of a novel protocadherin, OL-protocadherin, in a subset of functional systems of the developing mouse brain. Journal of Neuroscience 19:995–1005.

Horikawa K, Radice G, Takeichi M, Chisaka O (1999) Adhesive subdivisions intrinsic to the epithelial somites. Developmental biology. doi: 10.1006/dbio.1999.9463

Hovens CM, Stacker SA, Andres A-C, et al (1992) RYK, a Receptor Tyrosine Kinase-Related Molecule with Unusual Kinase Domain Motifs. Proceedings of the National Academy of Sciences of the United States of America 89:11818–11822.

Hubel, D.H., and Wiesel, T.N. (1963). Receptive Fields Of Cells In Striate Cortex Of Very Young, Visually Inexperienced Kittens. J. Neurophysiol. 26, 994–1002.

Hulpiau P, Gul IS, Van Roy F (2016) Evolution of Cadherins and Associated Catenins. The Cadherin Superfamily 13–37. doi: 10.1007/978-4-431-56033-3_2

Hulpiau P, Gul IS, Van Roy F (2013) New Insights into the Evolution of Metazoan Cadherins and Catenins. Progress in molecular biology and translational science 116:71–94. doi: 10.1016/B978-0-12-394311-8.00004-2

Hulpiau P, Van Roy F (2011) New insights into the evolution of metazoan cadherins. Molecular biology and evolution 28:647–657. doi: 10.1093/molbev/msq233

Hulpiau P, Van Roy F (2009) Molecular evolution of the cadherin superfamily. The International Journal of Biochemistry & Cell Biology 41:349–369. doi: 10.1016/j.biocel.2008.09.027

Hummel T, Vasconcelos ML, Clemens JC, et al (2003) Axonal Targeting of Olfactory Receptor Neurons in Drosophila Is Controlled by Dscam. Neuron 37:221–231. doi: 10.1016/S0896-6273(02)01183-2

Hussman JP, Chung R-H, Griswold AJ, et al (2011) A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. Molecular autism 2:1. doi: 10.1186/2040-2392-2-1

Huttenlocher PR (1979) Synaptic density in human frontal cortex - developmental changes and effects of aging. Brain research 163:195–205.

Hynes K, Tarpey P, Dibbens LM, et al (2010) Epilepsy and mental retardation limited to females with PCDH19 mutations can present de novo or in single generation families. Journal of Medical Genetics 47:211–216. doi: 10.1136/jmg.2009.068817

Ishii N, Wadsworth WG, Stern BD, et al (1992) UNC-6, a laminin-related protein, guides cell and pioneer axon migrations in *C. elegans*. Neuron 9:873–881.

Izuta Y, Taira T, Asayama A, et al (2015) Protocadherin-9 involvement in retinal development in Xenopus laevis. Journal of biochemistry 157:235–249. doi: 10.1093/jb/mvu070

Jorde, L.B., and Wooding, S.P. (2004). Genetic variation, classification and “race.” Nat. Genet. 36, S28–S33.

Juberg RC, Hellman CD (1971) A new familial form of convulsive disorder and mental retardation limited to females. The Journal of pediatrics 79:726–732.

Kai M, Ueno N, Kinoshita N (2015) Phosphorylation-Dependent Ubiquitination of Paraxial Protocadherin (PAPC) Controls Gastrulation Cell Movements. PLOS ONE 10:e0115111. doi: 10.1371/journal.pone.0115111

Kanai Y, Dohmae N, Hirokawa N (2004) Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. Neuron 43:513–525. doi: 10.1016/j.neuron.2004.07.022

Kaneko R, Kato H, Kawamura Y, et al (2006) Allelic gene regulation of Pcdh-alpha and Pcdh-gamma clusters involving both monoallelic and biallelic expression in single Purkinje cells. Journal of Biological Chemistry 281:30551–30560. doi: 10.1074/jbc.M605677200

Kang HJ, Kawasawa YI, Cheng F, et al (2011) Spatio-temporal transcriptome of the human brain. Nature 478:483–489. doi: 10.1038/nature10523

Katz MJ, Lasek RJ, Silver J (1983) Ontophyletics of the nervous system: development of the corpus callosum and evolution of axon tracts. Proceedings of the National Academy of Sciences 80:5936–5940.

Kayser MS, Nolt MJ, Dalva MB (2008) EphB receptors couple dendritic filopodia motility to synapse formation. Neuron 59:56–69. doi: 10.1016/j.neuron.2008.05.007

Keeble TR, Halford MM, Seaman C, et al (2006) The Wnt receptor Ryk is required for Wnt5a-mediated axon guidance on the contralateral side of the corpus callosum. The Journal of neuroscience : the official journal of the Society for Neuroscience 26:5840–5848. doi: 10.1523/JNEUROSCI.1175-06.2006

Kehayova P, Monahan K, Chen W, Maniatis T (2011) Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. Proceedings of the National Academy of Sciences of the United States of America 108:17195–17200. doi: 10.1073/pnas.1114357108

Keino-Masu K, Masu M, Hinck L, et al (1996) Deleted in Colorectal Cancer (DCC) Encodes a Netrin Receptor. Cell 87:175–185. doi: 10.1016/S0092-8674(00)81336-7

Kennedy TE, Serafini T, La Torre J De (1994) Netrins are diffusible chemotropic factors for commissural axons in the embryonic spinal cord. Cell 78:425–435. doi: 10.1016/0092-8674(94)90421-9

Kidd T, Bland KS, Goodman CS (1999) Slit Is the Midline Repellent for the Robo Receptor in Drosophila. Cell 96:785–794. doi: 10.1016/S0092-8674(00)80589-9

Kietzmann A, Wang Y, Weber D, Steinbeisser H (2012) Xenopus paraxial protocadherin inhibits Wnt/β-catenin signalling via casein kinase 2β. EMBO reports 13:129–134. doi: 10.1038/embor.2011.240

Kim SH, Yamamoto A, Bouwmeester T, et al (1998) The role of paraxial protocadherin in selective adhesion and cell movements of the mesoderm during Xenopus gastrulation. Development 125:4681–4690.

Kobayashi H, Koppel AM, Luo Y, Raper JA (1997) A role for collapsin-1 in olfactory and cranial sensory axon guidance. Journal of Neuroscience 17:8339–8352.

Koenig E (1967) Synthetic mechanisms in the axon. IV. In vitro incorporation of [3H]precursors into axonal protein and RNA. Journal of neurochemistry 14:437–446. doi: 10.1111/j.1471-4159.1967.tb09542.x

Kohmura N, Senzaki K, Hamada S, et al (1998) Diversity revealed by a novel family of cadherins expressed in neurons at a synaptic complex. Neuron 20:1137–1151.

Kolodkin AL, Matthes DJ, O'Connor TP, et al (1992) Fasciclin IV: Sequence, expression, and function during growth cone guidance in the grasshopper embryo. Neuron 9:831–845. doi: 10.1016/0896-6273(92)90237-8

Koning H, Sayers I, Stewart CE, et al (2012) Characterization of protocadherin-1 expression in primary bronchial epithelial cells: association with epithelial cell differentiation. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 26:439–448. doi: 10.1096/fj.11-185207

Koppelman GH, Meyers DA, Howard TD, et al (2009) Identification of PCDH1 as a novel susceptibility gene for bronchial hyperresponsiveness. American Journal of Respiratory and Critical Care Medicine 180:929–935. doi: 10.1164/rccm.200810-1621OC

Krishna-K K, Hertel N, Redies C (2011) Cadherin expression in the somatosensory cortex: evidence for a combinatorial molecular code at the single-cell level. 175:37–48. doi: 10.1016/j.neuroscience.2010.11.056

Krishna-K, Nuernberger M, Weth F, Redies C (2008) Layer-Specific Expression of Multiple Cadherins in the Developing Visual Cortex (V1) of the Ferret. 19:388–401. doi: 10.1093/cercor/bhn090

Kuroda H, Inui M, Sugimoto K, et al (2002) Axial Protocadherin Is a Mediator of Prenotochord Cell Sorting in Xenopus. Developmental biology 244:267–277. doi: 10.1006/dbio.2002.0589

La Manno G, Gyllborg D, Codeluppi S, et al (2016) Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. Cell 167:566–580. doi: 10.1016/j.cell.2016.09.027

Lee AC, Suter DM (2008) Quantitative analysis of microtubule dynamics during adhesion-mediated growth cone guidance. Developmental Neurobiology 68:1363–1377. doi: 10.1002/dneu.20662

Lefebvre JL, Kostadinov D, Chen WV, et al (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature 1–7. doi: 10.1038/nature11305

Leonardi E, Sartori S, Vecchi M, et al (2014) Identification of Four Novel PCDH19Mutations and Prediction of Their Functional Impact. Annals of Human Genetics 78:389–398. doi: 10.1111/ahg.12082

Lepeta K, Lourenco MV, Schweitzer BC, et al (2016) Synaptopathies: synaptic dysfunction in neurological disorders - A review from students to students. Journal of neurochemistry 138:785–805. doi: 10.1111/jnc.13713

Lescai F, Pirazzini C, D'Agostino G, et al (2010) Failure to replicate an association of rs5984894 SNP in the PCDH11X gene in a collection of 1,222 Alzheimer's disease affected patients. Journal of Alzheimer's disease : JAD 21:385–388. doi: 10.3233/JAD-2010-091516

Letourneau PC, Shattuck TA, Ressler AH (1987) 'Pull' and 'push' in neurite elongation: Observations on the effects of different concentrations of cytochalasin B and taxol. Cytoskeleton 8:193–209. doi: 10.1002/cm.970080302

Leung LC, Harris WA, Holt CE, Piper M (2015) NF-Protocadherin Regulates Retinal Ganglion Cell Axon Behaviour in the Developing Visual System. PLOS ONE 10:e0141290. doi: 10.1371/journal.pone.0141290

Leung LC, Urbančič V, Baudet M-L, et al (2013) Coupling of NF-protocadherin signaling to axon guidance by cue-induced translation. Nature neuroscience 16:166–173. doi: 10.1038/nn.3290

Levinson JN, Chéry N, Huang K, et al (2005) Neuroligins mediate excitatory and inhibitory synapse formation: involvement of PSD-95 and neurexin-1beta in neuroligin-induced synaptic specificity. Journal of Biological Chemistry 280:17312–17319. doi: 10.1074/jbc.M413812200

Levy S, Sutton G, Ng PC, et al (2007) The diploid genome sequence of an individual human. PLoS biology 5:e254. doi: 10.1371/journal.pbio.0050254

Lieberam I, Agalliu D, Nagasawa T, et al (2005) A Cxcl12-CXCR4 chemokine signaling pathway defines the initial trajectory of mammalian motor axons. Neuron 47:667–679. doi: 10.1016/j.neuron.2005.08.011

Lim BK, Matsuda N, Poo M (2008) Ephrin-B reverse signaling promotes structural and functional synaptic maturation in vivo. Nature neuroscience 11:160–169. doi: 10.1038/nn2033

Linask KK, Ludwig C, Han MD, et al (1998) N-cadherin/catenin-mediated morphoregulation of somite formation. Developmental biology 202:85–102. doi: 10.1006/dbio.1998.9025

Lin CH, Forscher P (1995) Growth cone advance is inversely proportional to retrograde F-actin flow. Neuron 14:763–771.

Lin DM, Auld VJ, Goodman CS (1995) Targeted neuronal cell ablation in the drosophila embryo: Pathfinding by follower growth cones in the absence of pioneers. Neuron 14:707–715. doi: 10.1016/0896-6273(95)90215-5

Lin J, Wang C, Redies C (2012) Expression of delta-protocadherins in the spinal cord of the chicken embryo. Journal of comparative neurology 520:1509–1531. doi: 10.1002/cne.22808

Lin Y-L, Ma J-H, Luo X-L, et al (2013) Clinical significance of protocadherin-8 (PCDH8) promoter methylation in bladder cancer. Journal of International Medical Research 41:48–54. doi: 10.1177/0300060513475571

Lin Y-L, Wang Y-L, Fu X-L, et al (2016) Low expression of protocadherin7 (PCDH7) is a potential prognostic biomarker for primary non-muscle invasive bladder cancer. Oncotarget 7:28384–28392. doi: 10.18632/oncotarget.8635

Liu J, Wilson S, Reh T (2003) BMP receptor 1b is required for axon guidance and cell survival in the developing retina. Developmental biology 256:34–48.

Liu Q, Chen Y, Pan JJ, Murakami T (2009) Expression of protocadherin-9 and protocadherin-17 in the nervous system of the embryonic zebrafish. Gene expression patterns : GEP 9:490–496. doi: 10.1016/j.gep.2009.07.006

Luckner R, Obst-Pernberg K, Hirano S, et al (2001) Granule cell raphes in the developing mouse cerebellum. Cell and tissue research 303:159–172. doi: 10.1007/s004410000292

Lumsden A, Davies AM (1983) Earliest sensory nerve fibres are guided to peripheral targets by attractants other than nerve growth factor. Nature. doi: 10.1038/306786a0

Luo R, Sanders SJ, Tian Y, et al (2012) Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. American journal of human genetics 91:38–55. doi: 10.1016/j.ajhg.2012.05.011

Luo Y, Raible D, Raper JA (1993) Collapsin: a protein in brain that induces the collapse and paralysis of neuronal growth cones. Cell. doi: 10.2307/25501340

Lu W, Yamamoto V, Ortega B, Baltimore D (2004) Mammalian Ryk Is a Wnt Coreceptor Required for Stimulation of Neurite Outgrowth. Cell 119:97–108. doi: 10.1016/j.cell.2004.09.019

MacNeil MA, Masland RH (1998) Extreme diversity among amacrine cells: implications for function. Neuron 20:971–982.

Makarenkova H, Sugiura H, Yamagata K, Owens G (2005) Alternatively spliced variants of protocadherin 8 exhibit distinct patterns of expression during mouse development. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression 1681:150–156. doi: 10.1016/j.bbaexp.2004.11.001

Malenka RC (1994) Synaptic plasticity in the hippocampus: LTP and LTD. Cell 78:535–538. doi: 10.1016/0092-8674(94)90517-7

Manser J, Wood WB (1990) Mutations affecting embryonic cell migrations in Caenorhabditis elegans. Developmental Genetics 11:49–64. doi: 10.1002/dvg.1020110107

Marini C, Darra F, Specchio N, et al (2012) Focal seizures with affective symptoms are a major feature of PCDH19gene-related epilepsy. Epilepsia 53:2111–2119. doi: 10.1111/j.1528-1167.2012.03649.x

Matsunaga M, Hatta K, Nagafuchi A, Takeichi M (1988) Guidance of optic nerve fibres by N-cadherin adhesion molecules. Nature 334:62–64. doi: 10.1038/334062a0

Mattar P, Britz O, Johannes C, et al (2004) A screen for downstream effectors of Neurogenin2 in the embryonic neocortex. Developmental biology 273:373–389.

Mayford M, Siegelbaum SA, Kandel ER (2012) Synapses and Memory Storage. Cold Spring Harbor perspectives in biology 4:a005751. doi: 10.1101/cshperspect.a005751

Medina A, Swain RK, Kuerner KM, Steinbeisser H (2004) Xenopus paraxial protocadherin has signaling functions and is involved in tissue separation. The EMBO journal 23:3249–3258. doi: 10.1038/sj.emboj.7600329

Miar A, Álvarez V, Corao AI, et al (2011) Lack of association between protocadherin 11-X/Y (PCDH11X and PCDH11Y) polymorphisms and late onset Alzheimer's disease. Brain research 1383:252–256. doi: 10.1016/j.brainres.2011.01.054

Miyatani S, Copeland NG, Gilbert DJ, et al (1992) Genomic structure and chromosomal mapping of the mouse N-cadherin gene. Proceedings of the National Academy of Sciences 89:8443–8447.

Molumby MJ, Anderson RM, Newbold DJ, et al (2017) γ-Protocadherins Interact with Neuroligin-1 and Negatively Regulate Dendritic Spine Morphogenesis. Cell Reports 18:2702–2714. doi: 10.1016/j.celrep.2017.02.060

Molumby MJ, Keeler AB, Weiner JA (2016) Homophilic Protocadherin Cell-Cell Interactions Promote Dendrite Complexity. Cell Reports 15:1037–1050. doi: 10.1016/j.celrep.2016.03.093

Morita A (2006) Regulation of Dendritic Branching and Spine Maturation by Semaphorin3A-Fyn Signaling. Journal of Neuroscience 26:2971–2980. doi: 10.1523/JNEUROSCI.5453-05.2006

Morrow EM, Yoo S-Y, Flavell SW, et al (2008) Identifying Autism Loci and Genes by Tracing Recent Shared Ancestry. Science 321:218–223. doi: 10.1126/science.1157657

Mortillo S, Elste A, Ge Y, et al (2012) Compensatory redistribution of neuroligins and N-cadherin following deletion of synaptic β1-integrin. Journal of comparative neurology 520:2041–2052. doi: 10.1002/cne.23027

Moscona A (1957). The Development in vitro of Chimeric Aggregates of Dissociated Embryonic Chick and Mouse Cells. Proc. Natl. Acad. Sci. U.S.A. 43, 184–194.

Moscona A, and Moscona H (1952). The dissociation and aggregation of cells from organ rudiments of the early chick embryo. J. Anat. 86, 287–301. doi: 10.1111/(ISSN)1469-7580

Mountoufaris G, Chen WV, Hirabayashi Y, et al (2017) Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. Science 356:411–414. doi: 10.1126/science.aai8801

Mulkey RM, Endo S, Shenolikar S, Malenka RC (1994) Involvement of a calcineurin/ inhibitor-1 phosphatase cascade in hippocampal long-term depression. Nature 369:486–488. doi: 10.1038/369486a0

Müller K, Hirano S, Puelles L, Redies C (2004) OL-protocadherin expression in the visual system of the chicken embryo. The Journal of comparative neurology 470:240–255. doi: 10.1002/cne.11044

Munton RP, Vizi S, Mansuy IM (2004) The role of protein phosphatase-1 in the modulation of synaptic and structural plasticity. FEBS letters 567:121–128. doi: 10.1016/j.febslet.2004.03.121

Murai KK, Nguyen LN, Irie F, et al (2003) Control of hippocampal dendritic spine morphology through ephrin-A3/EphA4 signaling. Nature neuroscience 6:153–160. doi: 10.1038/nn994

Myers JP, Santiago-Medina M, Gomez TM (2011) Regulation of axonal outgrowth and pathfinding by integrin-ECM interactions. Developmental Neurobiology 71:901–923. doi: 10.1002/dneu.20931

Nakao, S., Platek, A., Hirano, S., and Takeichi, M. (2008). Contact-dependent promotion of cell migration by the OL-protocadherin-Nap1 interaction. J. Cell Biol. 182, 395–410.

Nedivi E, Wu GY, Cline HT (1998) Promotion of dendritic growth by CPG15, an activity-induced signaling molecule. Science 281:1863–1866.

Neugebauer KM, Emmett CJ, Venstrom KA, Reichardt LF (1991) Vitronectin and thrombospondin promote retinal neurite outgrowth: developmental regulation and role of integrins. Neuron 6:345–358.

Neugebauer KM, Tomaselli KJ, Lilien J, Reichardt LF (1988) N-cadherin, NCAM, and integrins promote retinal neurite outgrowth on astrocytes in vitro. The Journal of cell biology 107:1177–1187.

Nicoludis JM, Vogt BE, Green AG, et al (2016) Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. eLife 5:213. doi: 10.7554/eLife.18449

Niewiadomska P, Godt D, Tepass U (1999) DE-Cadherin is required for intercellular motility during Drosophila oogenesis. The Journal of cell biology 144:533–547.

Nishimura, S.L., Boylen, K.P., Einheber, S., Milner, T.A., Ramos, D.M., and Pytela, R. (1998). Synaptic and glial localization of the integrin αvβ8 in mouse and rat brain. Brain Res. 791, 271–282.

Noda T, Satoh N (2008) A comprehensive survey of cadherin superfamily gene expression patterns in Ciona intestinalis. Gene Expression Patterns 8:349–356. doi: 10.1016/j.gep.2008.01.004

Nollet F, Kools P, Van Roy F (2000) Phylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. Journal of Molecular Biology 299:551–572. doi: 10.1006/jmbi.2000.3777

Nugent AA, Kolpak AL, Engle EC (2012) Human disorders of axon guidance. Current opinion in neurobiology 22:837–843. doi: 10.1016/j.conb.2012.02.006

Obata S, Sago H, Mori N, Davidson M (1998) A common protocadherin tail: multiple protocadherins share the same sequence in their cytoplasmic domains and are expressed in different regions of brain. Cell Communication … 6:323–333. doi: 10.3109/15419069809010791

Ohnami S, Endo M, Hirai S, et al (2008) Role of RhoA in activity-dependent cortical axon branching. The Journal of neuroscience : the official journal of the Society for Neuroscience 28:9117–9121. doi: 10.1523/JNEUROSCI.1731-08.2008

Okabe S, Miwa A, Okado H (2001) Spine formation and correlated assembly of presynaptic and postsynaptic molecules. The Journal of neuroscience : the official journal of the Society for Neuroscience 21:6105–6114.

Olivares R, Montiel J, Aboitiz F (2001) Species differences and similarities in the fine structure of the mammalian corpus callosum. Brain, behavior and evolution 57:98–105.

Omi M, Anderson R, Muneoka K (2002) Differential cell affinity and sorting of anterior and posterior cells during outgrowth of recombinant avian limb buds. Developmental biology 250:292–304.

Ostroff LE, Fiala JC, Allwardt B, Harris KM (2002) Polyribosomes redistribute from dendritic shafts into spines with enlarged synapses during LTP in developing rat hippocampal slices. Neuron 35:535–545.

Owen R (1837) On the Structure of the Brain in Marsupial Animals. Philosophical Transactions of the Royal Society of London 127:87–96.

Pakkenberg B (2003) Aging and the human neocortex. Experimental Gerontology 38:95–99. doi: 10.1016/S0531-5565(02)00151-1

Parker MW, Guo H-F, Li X, et al (2012) Function of Members of the Neuropilin Family as Essential Pleiotropic Cell Surface Receptors. Biochemistry 51:9437–9446. doi: 10.1021/bi3012143

Pederick DT, Homan CC, Jaehne EJ, et al (2016) Pcdh19 Loss-of-Function Increases Neuronal Migration In Vitro but is Dispensable for Brain Development in Mice. Nature Publishing Group 6:1–10. doi: 10.1038/srep26765

Pelvig DP, Pakkenberg H, Stark AK, Pakkenberg B (2008) Neocortical glial cell numbers in human brains. Neurobiology of Aging 29:1754–1762. doi: 10.1016/j.neurobiolaging.2007.04.013

Phillips KA, Stimpson CD, Smaers JB, et al (2015) The corpus callosum in primates: processing speed of axons and the evolution of hemispheric asymmetry. Proceedings Biological sciences 282:20151535. doi: 10.1098/rspb.2015.1535

Pini A (1993) Chemorepulsion of axons in the developing mammalian central nervous system. Science 261:95–98. doi: 10.1126/science.8316861

Pinkstaff JK, Detterich J, Lynch G, Gall C (1999) Integrin subunit gene expression is regionally differentiated in adult brain. Journal of Neuroscience 19:1541–1556.

Piper M, Dwivedy A, Leung L, et al (2008) NF-Protocadherin and TAF1 Regulate Retinal Axon Initiation and Elongation In Vivo. Journal of Neuroscience 28:100–105. doi: 10.1523/JNEUROSCI.4490-07.2008

Poliakov A, Cotrina M, Wilkinson DG (2004) Diverse Roles of Eph Receptors and Ephrins in the Regulation of Cell Migration and Tissue Assembly. Developmental cell 7:465–480. doi: 10.1016/j.devcel.2004.09.006

Poon VY, Klassen MP, Shen K (2008) UNC-6/netrin and its receptor UNC-5 locally exclude presynaptic components from dendrites. Nature 455:669–673. doi: 10.1038/nature07291

Price, S.R., De Marco Garcia, N.V., Ranscht, B., and Jessell, T.M. (2002). Regulation of motor neuron pool sorting by differential expression of type II cadherins. Cell 109, 205–216.

Priddle TH, Crow TJ (2013) Protocadherin 11X/Y a human-specific gene pair: an immunohistochemical survey of fetal and adult brains. Cerebral cortex (New York, NY : 1991) 23:1933–1941. doi: 10.1093/cercor/bhs181

Radice, GL, Rayburn, H, Matsunami, H, Knudsen, KA, Takeichi, M, and Hynes, RO (1997). Developmental defects in mouse embryos lacking N-cadherin. Developmental Biology 181, 64–78.

Rajasekharan S, Kennedy TE (2009) The netrin protein family. Genome Biology 10:239. doi: 10.1186/gb-2009-10-9-239

Ramón y Cajal S (1890) Notas anatómicas I. Sobre la aparición de las expansiones celulares en la médula embrionaria. Gac Sanit Barc 12:4139–4419.

Ramón y Cajal S (1892) La rétine des vertébrés. La Cellule 9:121–255.

Ramón y Cajal, S. (1928). Degeneration & regeneration of the nervous system (London: Oxford University Press, London).

Ravi V, Yu W-P, Pillai NE, et al (2016) Cyclostomes Lack Clustered Protocadherins. Molecular biology and evolution. Mol. Biol. Evol. 33:311–315. doi: 10.1093/molbev/msv252

Redies C, Heyder J, Kohoutek T, et al (2008) Expression of protocadherin-1 (Pcdh1) during mouse development. Developmental dynamics : an official publication of the American Association of Anatomists 237:2496–2505. doi: 10.1002/dvdy.21650

Redies, C., Vanhalst, K., and Roy, F.V. (2005). delta-Protocadherins: unique structures and functions. Cell. Mol. Life Sci. 62, 2840–2852.

Reijmers LG, Perkins BL, Matsuo N, Mayford M (2007) Localization of a Stable Neural Correlate of Associative Memory. Science 317:1230–1233. doi: 10.1126/science.1143839

Rhee J, Takahashi Y u, Saga Y, et al (2003) The protocadherin papc is involved in the organization of the epithelium along the segmental border during mouse somitogenesis. Developmental biology 254:248–261. doi: 10.1016/s0012-1606(02)00085-4

Riehl R, Johnson K, Bradley R, et al (1996) Cadherin function is required for axon outgrowth in retinal ganglion cells in vivo. Neuron 17:837–848.

Ross, N.L.J., Wadekar, R., Lopes, A., Dagnall, A., Close, J., DeLisi, L.E., and Crow, T.J. (2006). Methylation of two Homo sapiens-specific X-Y homologous genes in Klinefelter's syndrome (XXY). American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 141B, 544–548.

Roth SA, Weston JA (1967) The measurement of intercellular adhesion. Proceedings of the National Academy of Sciences 58:974–980.

Rubinstein R, Thu CA, Goodman KM, et al (2015) Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. Cell 163:1–15. doi: 10.1016/j.cell.2015.09.026

Rutishauser U, Gall WE, Edelman GM (1978a) Adhesion among neural cells of the chick embryo. IV. Role of the cell surface molecule CAM in the formation of neurite bundles in cultures of spinal ganglia. The Journal of cell biology 79:382–393. doi: 10.1083/jcb.79.2.382

Rutishauser U, Thiery JP, Brackenbury R, Edelman GM (1978b) Adhesion among neural cells of the chick embryo. III. Relationship of the surface molecule CAM to cell adhesion and the development of histotypic patterns. The Journal of cell biology 79:371–381. doi: 10.1083/jcb.79.2.371

Sahay A, Kim C-H, Sepkuty JP, et al (2005) Secreted semaphorins modulate synaptic transmission in the adult hippocampus. The Journal of neuroscience : the official journal of the Society for Neuroscience 25:3613–3620. doi: 10.1523/JNEUROSCI.5255-04.2005

Sano, K., Tanihara, H., Heimark, R.L., Obata, S., Davidson, M., St John, T., Taketani, S., and Suzuki, S. (1993). Protocadherins: a large family of cadherin-related molecules in central nervous system. Embo J. 12, 2249–2256.

Schmucker D, Clemens JC, Shu H, et al (2000) Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell 101:671–684. doi: 10.1016/S0092-8674(00)80878-8

Schneider M, Huang C, Becker SFS, et al (2014) Protocadherin PAPC is expressed in the CNC and can compensate for the loss of PCNS. Genesis (New York, NY : 2000) 52:120–126. doi: 10.1002/dvg.22736

Schoch H, Kreibich AS, Ferri SL, et al (2016) Sociability Deficits and Altered Amygdala Circuits in Mice Lacking Pcdh10, an Autism Associated Gene. Biological Psychiatry. doi: 10.1016/j.biopsych.2016.06.008

Schötz EM, Burdine RD, Jülicher F, et al (2008) Quantitative differences in tissue surface tension influence zebrafish germ layer positioning. HFSP Journal 2:42–56. doi: 10.2976/1.2834817

Schreiner D, Weiner JA (2010) Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. Proceedings of the National Academy of Sciences of the United States of America 107:14893–14898. doi: 10.1073/pnas.1004526107

Schwabe T, Borycz JA, Meinertzhagen IA (2014) Differential adhesion determines the organization of synaptic fascicles in the Drosophila visual system. Current biology 24:1304–1313. doi: 10.1016/j.cub.2014.04.047

Serafini T, Kennedy TE, Galko MJ, et al (1994) The netrins define a family of axon outgrowth-promoting proteins homologous to *C. elegans* UNC-6. Cell 78:409–424.

Serizawa S, Miyamichi K, Takeuchi H, et al (2006) A neuronal identity code for the odorant receptor-specific and activity-dependent axon sorting. Cell 127:1057–1069. doi: 10.1016/j.cell.2006.10.031

Shirayoshi Y, Okada TS, Takeichi M (1983) The calcium-dependent cell-cell adhesion system regulates inner cell mass formation and cell surface polarization in early mouse development. Cell 35:631–638. doi: 10.1016/0092-8674(83)90095-8

Smith GE (1910) The Arris and Gale Lectures On Some Problems Relating To The Evolution Of The Brain. The Lancet 175:221–227. doi: 10.1016/S0140-6736(01)14322-9

Song JY, Ichtchenko K, Südhof TC, Brose N (1999) Neuroligin 1 is a postsynaptic cell-adhesion molecule of excitatory synapses. Proceedings of the National Academy of Sciences 96:1100–1105.

Sotelo C (2002) The chemotactic hypothesis of Cajal: a century behind. Progress in brain research 136:11–20. doi: 10.1016/S0079-6123(02)36004-7

Specchio, N., Fusco, L., and Vigevano, F. (2011). Acute-onset epilepsy triggered by fever mimicking FIRES (febrile infection-related epilepsy syndrome): the role of protocadherin 19 (PCDH19) gene mutation. Epilepsia 52, e172–e175.

Speevak MD, Farrell SA (2011) Non-syndromic language delay in a child with disruption in the Protocadherin11X/Y gene pair. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 156B:484–489. doi: 10.1002/ajmg.b.31186

Spemann H, Mangold H (1924) über Induktion von Embryonalanlagen durch Implantation artfremder Organisatoren. Archiv für Mikroskopische Anatomie und Entwicklungsmechanik 100:599–638. doi: 10.1007/BF02108133

Stagi M, Fogel AI, Biederer T (2010) SynCAM 1 participates in axo-dendritic contact assembly and shapes neuronal growth cones. Proceedings of the National Academy of Sciences 107:7568–7573. doi: 10.1073/pnas.0911798107

Steinberg, MS (1958). On the Chemical Bonds between Animal Cells. A Mechanism for Type-Specific Association. The American Naturalist *92*, 65–81.

Steinberg, MS (1963). “ECM”: its nature, origin and function in cell aggregation. Exp. Cell Res. *30*, 257–279.

Steinberg, MS, and Garrod, D.R. (1975). Observations on the sorting-out of embryonic cells in monolayer culture. Journal of Cell Science *18*, 385–403.

Steinberg, MS (2007). Differential adhesion in morphogenesis: a modern view. Curr. Opin. Genet. Dev. *17*, 281–286.

Steinberg, MS, and Takeichi, M. (1994). Experimental specification of cell sorting, tissue spreading, and specific spatial patterning by quantitative differences in cadherin expression. Pnas *91*, 206–209.

Steward O, Levy WB (1982) Preferential localization of polyribosomes under the base of dendritic spines in granule cells of the dentate gyrus. Journal of Neuroscience 2:284–291.

Steward O, Ribak CE (1986) Polyribosomes associated with synaptic specializations on axon initial segments: localization of protein-synthetic machinery at inhibitory synapses. Journal of Neuroscience 6:3079–3085.

Stiles, J., and Jernigan, T.L. (2010). The basics of brain development. Neuropsychol Rev 20, 327–348.

Strehl S, Glatt K, Liu QM, et al (1998) Characterization of two novel protocadherins (PCDH8 and PCDH9) localized on human chromosome 13 and mouse chromosome 14. Genomics 53:81–89. doi: 10.1006/geno.1998.5467

Sugino H, Hamada S, Yasuda R, et al (2000) Genomic Organization of the Family of CNR Cadherin Genes in Mice and Humans. Genomics 63:75–87. doi: 10.1006/geno.1999.6066

Suo, L., Lu, H., Ying, G., Capecchi, M.R., and Wu, Q. (2012). Protocadherin clusters and cell adhesion kinase regulate dendrite complexity through Rho GTPase. J Mol Cell Biol 4, 362–376.

Szabó N-E, Haddad-Tóvolli R, Zhou X, Alvarez-Bolado G (2015) Cadherins mediate sequential roles through a hierarchy of mechanisms in the developing mammillary body. Frontiers in neuroanatomy 9:39. doi: 10.3389/fnana.2015.00029

Tai K, Kubota M, Shiono K, Tokutsu H, and Suzuki S (2010). Adhesion properties and retinofugal expression of chicken protocadherin-19. Brain Res. 1344, 13–24.

Takeichi M (1977) Functional correlation between cell adhesive properties and some cell surface proteins. The Journal of cell biology 75:464–474.

Takeichi, M., Hatta, K., Nose, A., and Nagafuchi, A. (1988). Identification of a gene family of cadherin cell adhesion molecules. Cell Differ. Dev. 25 Suppl, 91–94.

Tamagnone L, Artigiani S, Chen H, et al (1999) Plexins Are a Large Family of Receptors for Transmembrane, Secreted, and GPI-Anchored Semaphorins in Vertebrates. Cell 99:71–80. doi: 10.1016/S0092-8674(00)80063-X

Tan C, Shard C, Ranieri E, Hynes K, Pham DH, Leach D, Buchanan G, Corbett M, Shoubridge C, Kumar R, et al. (2015). Mutations of protocadherin 19 in female epilepsy (PCDH19-FE) lead to allopregnanolone deficiency. Hum. Mol. Genet. 24, 5250–5259.

Tasic B, Menon V, Nguyen TN, et al (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature neuroscience 19:335–346. doi: 10.1038/nn.4216

Tasic B, Nabholz CE, Baldwin KK, et al (2002) Promoter choice determines splice site selection in protocadherin alpha and gamma pre-mRNA splicing. Molecular Cell 10:21–33.

Tessier-Lavigne M, Placzek M, Lumsden AG, et al (1988) Chemotropic guidance of developing axons in the mammalian central nervous system. Nature 336:775–778. doi: 10.1038/336775a0

Thiery JP, Brackenbury R, Rutishauser U, Edelman GM (1977) Adhesion among neural cells of the chick embryo. II. Purification and characterization of a cell adhesion molecule from neural retina. Journal of Biological Chemistry 252:6841–6845.

Thu CA, Chen WV, Rubinstein R, et al (2014) Single-cell identity generated by combinatorial homophilic interactions between α, β, and γ protocadherins. Cell 158:1045–1059. doi: 10.1016/j.cell.2014.07.012

Togashi H, Abe K, Mizoguchi A, et al (2002) Cadherin regulates dendritic spine morphogenesis. Neuron 35:77–89.

Toncheva AA, Suttner K, Michel S, et al (2012) Genetic variants in Protocadherin-1, bronchial hyper-responsiveness, and asthma subphenotypes in German children. Pediatric Allergy and Immunology 23:636–641. doi: 10.1111/j.1399-3038.2012.01334.x

Townes, P.L., and Holtfreter, J. (1955). Directed movements and selective adhesion of embryonic amphibian cells. Journal of Experimental Zoology 128:53–120.

Toyoda S, Kawaguchi M, Kobayashi T, et al (2014) Developmental epigenetic modification regulates stochastic expression of clustered protocadherin genes, generating single neuron diversity. Neuron 82:94–108. doi: 10.1016/j.neuron.2014.02.005

Trachtenberg JT, Chen BE, Knott GW, Feng G (2002) Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. Nature 420:788–794. doi: 10.1038/nature01273

Trinkaus JP, Groves PW (1955) Differentiation In Culture Of Mixed Aggregates Of Dissociated Tissue Cells. Proceedings of the National Academy of Sciences 41:787–795.

Trivisano M, Lucchi C, Rustichelli C, et al (2017) Reduced steroidogenesis in patients with PCDH19-female limited epilepsy. Epilepsia 13:35. doi: 10.1111/epi.13772

Trivisano M, Pietrafusa N, Ciommo V di, et al (2016) PCDH19-related epilepsy and Dravet Syndrome: Face-off between two early-onset epilepsies with fever sensitivity. Epilepsy research 125:32–36. doi: 10.1016/j.eplepsyres.2016.05.015

Trousse F, Martí E, Gruss P, et al (2001) Control of retinal ganglion cell axon growth: a new role for Sonic hedgehog. Development 128:3927–3936.

Tsai N-P, Wilkerson JR, Guo W, et al (2012) Multiple autism-linked genes mediate synapse elimination via proteasomal degradation of a synaptic scaffold PSD-95. Cell 151:1581–1594. doi: 10.1016/j.cell.2012.11.040

Uemura M, Nakao S, Suzuki ST, et al (2007) OL-Protocadherin is essential for growth of striatal axons and thalamocortical projections. 10:1151–1159. doi: 10.1038/nn1960

Urushihara H, Ozaki HS, Takeichi M (1979) Immunological detection of cell surface components related with aggregation of Chinese hamster and chick embryonic cells. Developmental biology 70:206–216. doi: 10.1016/0012-1606(79)90017-4

Urushihara H, Takeichi M (1980) Cell-cell adhesion molecule: identification of a glycoprotein relevant to the Ca2+-independent aggregation of Chinese hamster fibroblasts. Cell 20:363–371.

Usoskin D, Furlan A, Islam S, et al (2015) Unbiased classification of sensory neuron types by large-scale single-cell RNA sequencing. Nature neuroscience 18:145–153. doi: 10.1038/nn.3881

Van Battum MSc EY, MSc SB, PhD DRJP (2015) Axon guidance proteins in neurological disorders. The Lancet Neurology 14:532–546. doi: 10.1016/S1474-4422(14)70257-1

Vanhalst, K., Kools, P., Staes, K., van Roy, F., and Redies, C. (2005). delta-Protocadherins: a gene family expressed differentially in the mouse brain. Cell. Mol. Life Sci. 62, 1247–1259.

Veerappa AM, Saldanha M, Padakannaya P, Ramachandra NB (2013) Genome-wide copy number scan identifies disruption of PCDH11X in developmental dyslexia. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 162:889–897. doi: 10.1002/ajmg.b.32199

Wang J, Zugates CT, Liang IH, et al (2002a) Drosophila Dscam Is Required for Divergent Segregation of Sister Branches and Suppresses Ectopic Bifurcation of Axons. Neuron 33:559–571. doi: 10.1016/S0896-6273(02)00570-6

Wang X, Weiner JA, Levi S, et al (2002b) Gamma protocadherins are required for survival of spinal interneurons. Neuron 36:843–854.

Wang Y, Janicki P, Köster I, et al (2008) Xenopus Paraxial Protocadherin regulates morphogenesis by antagonizing Sprouty. Genes & development 22:878–883. doi: 10.1101/gad.452908

Washbourne P, Bennett JE, McAllister AK (2002) Rapid recruitment of NMDA receptor transport packets to nascent synapses. Nature neuroscience 5:751–759. doi: 10.1038/nn883

Watari-Goshima N, Ogura K, Wolf FW, et al (2007) *C. elegans* VAB-8 and UNC-73 regulate the SAX-3 receptor to direct cell and growth-cone migrations. Nature neuroscience 10:169–176. doi: 10.1038/nn1834

Weiner, JA, and Jontes, JD (2013). Protocadherins, not prototypical: a complex tale of their interactions, expression, and functions. Front Mol Neurosci 6, 4.

Weiner JA, Wang X, Tapia JC, Sanes JR (2005) Gamma protocadherins are required for synaptic development in the spinal cord. Proceedings of the National Academy of Sciences of the United States of America 102:8–14. doi: 10.1073/pnas.0407931101

Wiesel TN, Hubel DH (1963a) Single-Cell Responses In Striate Cortex Of Kittens Deprived Of Vision In One Eye. Journal of neurophysiology 26:1003–1017.

Wiesel, TN, and Hubel, DH (1963b). Effects Of Visual Deprivation On Morphology And Physiology Of Cells In The Cats Lateral Geniculate Body. J. Neurophysiol. 26, 978–993.

Wightman B, Clark SG, Taskar AM, et al (1996) The *C. elegans* gene vab-8 guides posteriorly directed axon outgrowth and cell migration. Development 122:671–682.

Williams EO, Sickles HM, Dooley AL, et al (2011) Delta Protocadherin 10 is Regulated by Activity in the Mouse Main Olfactory System. Frontiers in Neural Circuits. doi: 10.3389/fncir.2011.00009

Wilson HV (1907) On some phenomena of coalescence and regeneration in sponges. Journal of Experimental Zoology 5:245–258. doi: 10.1002/jez.1400050204

Wojtowicz WM, Flanagan JJ, Millard SS, et al (2004) Alternative Splicing of Drosophila Dscam Generates Axon Guidance Receptors that Exhibit Isoform-Specific Homophilic Binding. Cell 118:619–633. doi: 10.1016/j.cell.2004.08.021

Wojtowicz WM, Wu W, Andre I, et al (2007) A Vast Repertoire of Dscam Binding Specificities Arises from Modular Interactions of Variable Ig Domains. Cell 130:1134–1145. doi: 10.1016/j.cell.2007.08.026

Wolf FW, Hung M-S, Wightman B, et al (1998) vab-8 Is a Key Regulator of Posteriorly Directed Migrations in *C. elegans* and Encodes a Novel Protein with Kinesin Motor Similarity. Neuron 20:655–666. doi: 10.1016/S0896-6273(00)81006-5

Wolverton T, Lalande M (2001) Identification and characterization of three members of a novel subclass of protocadherins. Genomics 76:66–72. doi: 10.1006/geno.2001.6592

Wu C, Niu L, Yan Z, et al (2015) Pcdh11x Negatively Regulates Dendritic Branching. Journal of Molecular Neuroscience 56:822–828. doi: 10.1007/s12031-015-0515-8

Wu Q, Maniatis T (2000) Large exons encoding multiple ectodomains are a characteristic feature of protocadherin genes. Proceedings of the National Academy of Sciences of the United States of America 97:3124–3129. doi: 10.1073/pnas.060027397

Wu Q, Zhang T, Cheng JF, et al (2001) Comparative DNA sequence analysis of mouse and human protocadherin gene clusters. Genome research 11:389–404. doi: 10.1101/gr.167301

Xiang Y, Li Y, Zhang Z, et al (2002) Nerve growth cone guidance mediated by G protein-coupled receptors. Nature neuroscience 5:843–848. doi: 10.1038/nn899

Yagi T (2012) Molecular codes for neuronal individuality and cell assembly in the brain. Frontiers in molecular neuroscience 5:45. doi: 10.3389/fnmol.2012.00045

Yajima H, Tamura SY, and Watanabe N (1999). Role of N-cadherin in the sorting-out of mesenchymal cells and in the positional identity along the proximodistal axis of the chick limb bud. Developmental Dynamics 216, 274–284.

Yamagata K, Andreasson KI, Sugiura H, et al (1999) Arcadlin is a neural activity-regulated cadherin involved in long term potentiation. Journal of Biological Chemistry 274:19473.

Yamakawa K, Huot YK, Haendelt MA, Hubert R, Chen XN, Lyons GE, and Korenberg JR (1998). DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. Hum. Mol. Genet. *7*, 227–237.

Yamamoto A, Kemp C, Bachiller D, et al (2000) Mouse paraxial protocadherin is expressed in trunk mesoderm and is not essential for mouse development. Genesis (New York, NY : 2000) 27:49–57.

Yamashita N, Morita A, Uchida Y, et al (2007) Regulation of Spine Development by Semaphorin3A through Cyclin-Dependent Kinase 5 Phosphorylation of Collapsin Response Mediator Protein 1. Journal of Neuroscience 27:12546–12554. doi: 10.1523/JNEUROSCI.3463-07.2007

Yamauchi K, Phan KD, Butler SJ (2008) BMP type I receptor complexes have distinct activities mediating cell fate and axon guidance decisions. Development 135:1119–1128. doi: 10.1242/dev.012989

Yasuda S, Tanaka H, Sugiura H, et al (2007) Activity-Induced Protocadherin Arcadlin Regulates Dendritic Spine Number by Triggering N-Cadherin Endocytosis via TAO2β and p38 MAP Kinases. Neuron 56:456–471. doi: 10.1016/j.neuron.2007.08.020

Yoder MD, Gumbiner BM (2011) Axial protocadherin (AXPC) regulates cell fate during notochordal morphogenesis. Developmental Dynamics 240:2495–2504. doi: 10.1002/dvdy.22754

Yoshida K (2003) Fibroblast cell shape and adhesion in vitro is altered by overexpression of the 7a and 7b isoforms of protocadherin 7, but not the 7c isoform. Cellular & molecular biology letters 8:735–741.

Yoshida K, Watanabe M, Kato H, et al (1999) BH-protocadherin-c, a member of the cadherin superfamily, interacts with protein phosphatase 1 alpha through its intracellular domain. FEBS letters 460:93–98. doi: 10.1016/S0014-5793(99)01309-5

Yoshida K, Yoshitomo-Nakagawa K, Seki N, et al (1998) Cloning, Expression Analysis, and Chromosomal Localization of BH-Protocadherin (PCDH7), a Novel Member of the Cadherin Superfamily. Genomics 49:458–461. doi: 10.1006/geno.1998.5271

Yoshikawa S, Bonkowsky JL, Kokel M (2001) The derailed guidance receptor does not require kinase activity in vivo. The Journal of Neuroscience 21:RC119 1-4.

Yoshimura Y, Yamauchi Y, Shinkawa T, et al (2004) Molecular constituents of the postsynaptic density fraction revealed by proteomic analysis using multidimensional liquid chromatography-tandem mass spectrometry. Journal of neurochemistry 88:759–768. doi: 10.1046/j.1471-4159.2003.02136.x

Yu Q, Li X-T, Liu C, et al (2014) Evolutionarily conserved repulsive guidance role of slit in the silkworm Bombyx mori. PLOS ONE 9:e109377. doi: 10.1371/journal.pone.0109377

Zeisel A, Muñoz-Manchado AB, Codeluppi S, et al (2015) Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347:1138–1142. doi: 10.1126/science.aaa1934

Zhai RG, Vardinon-Friedman H, Cases-Langhoff C, et al (2001) Assembling the presynaptic active zone: a characterization of an active one precursor vesicle. Neuron 29:131–143. doi: 10.1016/S0896-6273(01)00185-4

Zhang P, Wu C, Liu N, et al (2014) Protocadherin 11 X Regulates Differentiation and Proliferation of Neural Stem Cell In Vitro and In Vivo. Journal of Molecular Neuroscience 54:199–210. doi: 10.1007/s12031-014-0275-x

Zhou X, Updegraff BL, Guo Y, et al (2017) Protocadherin 7 Acts through SET and PP2A to Potentiate MAPK Signaling by EGFR and KRAS during Lung Tumorigenesis. Cancer Research 77:187–197. doi: 10.1158/0008-5472.CAN-16-1267-T

Chapter 2

Defining Combinatorial Expression of δ-Protocadherins in Neurons Using Single-Cell Analysis

1. Abstract

Understanding how neural wiring is specified in the mammalian brain will greatly enhance our understanding of brain function and dysfunction. The daunting task of instructing neural wiring choices has long been proposed to utilize combinations of molecular cues. However, to understand how combinations may function in neural wiring, it is critical to first know what they are, and how complex they may be. Defining these combinations has been difficult due to the vast complexity of the nervous system and its great deal of heterogeneity. Now, with the capability to study gene expression in single cells, the combinatorial nature of these cues can begin to be revealed so that their functions may be further studied to understand how they regulate neural wiring. Here, we develop a targeted single-cell gene expression method to detect ~35 genes in 50 single olfactory sensory neurons using the NanoString platform. We focus on the δ-protocadherins, a small cell-adhesion family that has been implicated in neural development and proposed to function via combinatorial expression in neurons. Here, we report that the number of δ-protocadherins that can be expressed in individual olfactory sensory neurons ranges from zero to seven, and we validated our findings

using two independent RNA detection methods. Finally, we compare our results to single-cell RNA sequencing (scRNA-seq) studies and find that our targeted method can detect more δ-protocadherins per neuron than scRNA-seq studies.

2. Introduction

A longstanding hypothesis of neural development is that neurons utilize combinations of molecular cues to specify neural wiring choices. This has been difficult to study, however, because the nervous system is immensely complex due to the vast number of cells it contains, a high degree of cellular heterogeneity, and the trillions of synaptic connections made between them. While the function of the nervous system depends on the precise arrangement of neural circuits, understanding how these circuits are formed has been technically and conceptually challenging because neurons appear to express different combinations of numerous genes that control for axon guidance and synaptogenesis. Even though some studies have begun to demonstrate why combinations are important (Stevens and Jacobs 2002; Rhee et al. 2007; Emond et al. 2011), it has been difficult to understand what the combinations are, how many genes are expressed, and how different they may be among neurons. Now, technological advancements have enabled us to begin answering these questions. The ability to profile the nucleic acids of single cells has given researchers unprecedented insight into the large amount of complex heterogeneity that exists in biology. Even though we are still in the infancy of its technological development, single-cell analyses have already revealed many important discoveries about biological complexity that could not be

previously addressed. For example, single-cell studies have uncovered the molecular changes during developmental pathways (Tan et al. 2015; Hanchate et al. 2015; Petropoulos et al. 2016) and provided new understandings about intra-tumoral heterogeneity (Patel et al. 2014; Chung et al. 2017) and the clonal evolution of tumors (Navin et al. 2011; Wang et al. 2014; Ross and Markowetz 2016).

Single-cell analyses can therefore greatly enhance our understanding of the nervous system by defining the molecular diversity of neurons to begin untangling its complexity. Initial efforts to study gene expression in single neurons relied primarily on PCR-based approaches such as RT-PCR (Sucher et al. 2000; Lin et al. 2007; Toledo-Rodriguez and Markram 2007), and allowed researchers to examine a small number of genes of interest in a non-quantitative manner (Mackler and Eberwine 1993; Miyashiro et al. 1994; Crino et al. 1996; Baro et al. 1997; Horikoshi and Sakakibara 2000; Esumi et al. 2005; Esumi et al. 2006; Hirano et al. 2012). The use of qRT-PCR increased throughout slightly and enabled quantification (Citri et al. 2011) and microarrays vastly enhanced the number of genes that could be studied (Tietjen et al. 2003; Kamme et al. 2003; Tietjen et al. 2005). However, since these methods are highly PCR dependent but have no means for correcting or controlling amplification artifacts, they can suffer from strong amplification biases, sensitivity limitations, and technical noise.

Today, the most common approach is single-cell RNA-seq (scRNA-seq). First demonstrated in 2009 (Tang et al. 2009), many variations for profiling the transcriptomes of individual cells with scRNA-seq have since been developed (Islam et al. 2011; Ramsköld et al. 2012; Picelli et al. 2013; Macosko et al. 2015; Gierahn et al.

2017), each with their own advantages and disadvantages (Kolodziejczyk et al. 2015; Svensson et al. 2017). The high-throughput capabilities of transcriptome profiling has already allowed many sub-populations of neurons to be identified, increasing our knowledge of the cellular diversity in the nervous system (Tasic et al. 2016; Shekhar et al. 2016; La Manno et al. 2016). However, although many technical aspects of scRNA-seq have been improved such as the use of unique molecular identifiers (UMIs) to remove amplification biases (Kivioja et al. 2012; Islam et al. 2014), scRNA-seq studies suffer from two major limitations for understanding the relationship between single neuron gene expression, and neuronal function.

First, a critical restriction with scRNA-seq is the low capture efficiency of mRNA molecules within a cell. The amount of total RNA in single diploid mammalian cells has been measured to be in the range of 20-40 pg, depending on the cell type (Edström 1964; Roozemond 1976; Uemura 1980). The proportion of that which is mRNA is estimated to be minimal, comprising of only 1-2%, creating the most challenging aspect of single-cell analyses (Kawasaki 2004). Problematically, current scRNA-seq methods are estimated to have very poor capture efficiency of the mRNA in a cell, with studies estimating only 10-50% is captured per cell (Islam et al. 2014; Marinov et al. 2014). Since it is known that some genes are only expressed in low copies in individual cells (Islam et al. 2011), a significant proportion of genes and transcripts will inevitably be undetected or under represented. Remarkably, this appears to not significantly hinder general classification of cell types via hierarchical clustering. However, missing a large portion of a cell's mRNA will significantly skew any biological interpretations that are

focused on particular genes. For example, if you are attempting to define combinations of genes, low capture efficiency could cause misleading results.

In this study, we developed a medium-throughput single-cell analysis method. Instead of attempting to capture and sequence the entire transcriptome, our approach targeted 35 specific genes involved in neural wiring and used the NanoString nCounter digital hybridization platform (Geiss et al. 2008; Malkov et al. 2009) for detection. Using this method, we analyzed 50 randomly selected olfactory sensory neurons (OSNs). We focused on the δ-protocadherins, a family of nine cell-adhesion molecules (Vanhalst et al. 2005; Redies et al. 2005) that have been implicated in various aspects of neural development (Yamagata et al. 1999; Uemura et al. 2007; Yasuda et al. 2007; Bononi et al. 2008; Emond et al. 2009; Schneider et al. 2014; Wu et al. 2015) and neurological conditions (Dibbens et al. 2008; Morrow et al. 2008; Bucan et al. 2009; Luo et al. 2012; Tsai et al. 2012; Butler et al. 2015). Expression studies throughout the nervous system have led to the proposal that they are combinatorially expressed in individual neurons (Vanhalst et al. 2005; Hertel et al. 2008; Etzrodt et al. 2009; Krishna-K et al. 2011), and that this would generate a “combinatorial code” of adhesive specificities for neural development and neural circuit formation. However, this has never been properly examined or defined.

The purpose of this study was to determine if the δ-protocadherins are combinatorially expressed in single OSNs, and if so, to what extent. Our results show that in individual OSNs, the δ-protocadherins are expressed in a wide range of combinations, and we validate our findings using RNA *in situ* hybridizations and single-cell qPCR. We also

show that our targeted method appears to be better at detecting δ-protocadherins in single neurons than scRNA-seq by comparing our results to publicly available data sets. By defining the depth of combinatorial expression that exists in neurons, we have now set the stage to begin to study δ-protocadherin function in the context of their combinatorial possibilities.

3. Methods

3.1. Animal Use

All protocols for treatment and breeding of animals were approved by the Cornell Institutional Animal Care Use Committee (IACUC). Non-Swiss Albino (NSA) mice were used for all single-cell studies. For *in situ* RNA hybridization experiments, both NSA and C57Bl/6 mice were used. For all single-cell and single label *in situ* experiments, mice were sacrificed at post-natal day 7 (P7).

3.2. *In Situ* RNA Hybridizations and Analysis

P7 mice were euthanized and decapitated. Heads were then embedded in OCT (TissueTek) and fresh-frozen on isopentane cooled by liquid nitrogen. 20 μM cryosections were collected and allowed to dry for 1-2 hours, fixed with 4% paraformaldehyde for 15 minutes, and then washed three times in PBS. The slides were then acetylated in 0.25% acetic anhydride in 0.1M triethanolamine for 15 minutes, and washed again three times in PBS. Slides were then blocked in hybridization buffer for several hours at room temperature. Digoxigenin labeled antisense RNA probes were

prepared at 1ng/uL in hybridization buffer, heated at 80C for 10 minutes, and then cooled on ice for 3 minutes. All probes targeted the nucleotides consisting of the whole ectodomain and transmembrane domain of the δ-protocadherins. Probes were then allowed to hybridized with the tissue for 24-48 hours at 65-72C depending on the probe. After hybridization, slides were washed in 2X SSC at hybridization temperature for 30 minutes, three times with 0.2X SSC at hybridization temperature for 15 minutes, and then once with 0.2X SSC at room temperature. The slides were then washed with B1 (100 mM Tris pH 7.4, 150 mm NaCl) for 15 minutes and then blocked (Roche Blocking reagent) for ~2 hours at room temperature before incubating with anti-digoxigenin-alkaline phosphatase antibody at 1:3,000 overnight at 4C. The next day, the slides were wash three times with B1 + 0.05% Tween 20 for 15 minutes each, and then with B3 (100 mM Tris pH 9.5, 50 mM $MgCl_2$, 100 mM NaCl) for 15 minutes. Finally, the slides were reacted with NBT/BCIP (Promega, 15 μL NBT + 7.2 μ BCIP per 1.5 mL of B3). Replicates from three different heads were performed for each gene, and all slides were scanned using an Aperio Scanscope. To quantify percent area expressed in OSNs, each neural layer of the epithelium was manually traced and analyzed using a Halo imaging algorithm (Indica Labs).

3.3. Single OSN Isolation

The olfactory epithelium of P7 NSA mice was dissected and enzymatically dissociated into single cell suspensions. Approximately 250,000 cells were then plated on glass coverslips coated with poly-ornithine, and the cells were allowed to recover at 37C with 6% CO_2 for 30 minutes in Modified Eagle's Medium (MEM). After recovering, the cells

were gently washed with incubator equilibrated MEM three times. The coverslip was then transferred into a 10 cm dish, immobilized by applying small dabs of autoclaved Vaseline between the bottom of the coverslip and the 10 cm dish, and then adding 10 mL of incubator equilibrated MEM. Individual cells were then isolated using manual aspiration with micropipettes under a 20X objective. Micropipettes were prepared using a Sutter P-97 Flaming/Brown micropipette puller. Tips were pulled such that the bore of the pipette was only slightly larger than the diameter of an OSN. The following pulling parameters were used: Heat = 600, Pull = 7, Time = 200, Velocity = 5, Pressure = 620. Prior to isolation, each pipette was filled with ~3 µL of MEM. While a neuron was being selected for isolation, positive pressure was slowly applied to the micropipette to prevent media and any contaminants from entering the micropipette before the chosen neuron was isolated. Two different single cell lysis/RNA kits were used during cell collections, Cells Direct (Invitrogen) and Cells-to-Ct (Ambion). In house control tests showed that they did not produce significantly different results. After an OSN was aspirated into the micropipette, the contents (usually 1-2 µL) were transferred into a PCR tube containing 6 µL of lysis buffer mix (for Ambion, 9uL lysis + 1 µL DNAseI; for Cells Direct, 10 µL lysis buffer + 1 µL enhancer) by gently snapping the most distal tip of the micropipette inside the tube at a 45 degree angle to increase the bore size, and then expelling the contents using an intramuscular needle and syringe. To minimize time out of the incubator, neurons were isolated from each coverslip for no more than 30 minutes. If cells began to look unhealthy, cell picking was stopped. Cells processed in Cells Direct lysis buffer could immediately be stored at -80 until ready for processing.

For Ambion samples, cell lysates were vortexed, and incubated at room temperature for 5 minutes, and then 0.5 µL of the stop solution was added and incubated for 2 minutes at room temperature before being stored at -80C until further processing.

3.4. Processing and Screening of Single-Cell RNA

To process the single-cell RNA into cDNA, individual cell lysates were brought up to 10 µL with DEPC treated and the 35 gene multiplex primer set (100 nM final, see Table 2.1), and heated at 80C for 10 minutes and then chilled on ice for 3 minutes. Next, 10uL of 2x buffer and 1 µL of the reverse transcriptase III/PlatinumTaq mix were added and incubated at 50C for 1 hour. Immediately after, the samples were then heated to 85C for 15 minutes, and then 94C for 2 minutes, followed by 18 cycles of 94C for 30 seconds, 60C for 30 seconds, 72C for 30 seconds.

After amplification, 20 µL of 10 mM Tris 7.5 was added to each sample to bring up the volume to 40 and then stored at -80C until further analysis. Each processed neuron was then screened using 2 μL of its cDNA as a template for nested Taqman qPCR reactions for the marker genes Gapdh and Ncam1. Samples were run on an ABI 7500 under standard cycling conditions, and Ct values were determined using AutoCt detection. Only cells that had Ct values ≤ 25 for both genes were used for the NanoString analysis. The following primer sequences were used: Gapdh (forward): AACAGCAACTCCCACTCTTC, Gapdh (reverse): AATACGGCTACAGCAACAGG, Gapdh (probe): GGCTGGCATTGCTCTCAATGACAA, Ncam1 (forward): GATGCTGTGATTGTCTGTGAT, Ncam1 (reverse): GACTATGAACCGGACGTCTTT,

Ncam1 (probe): CCTGCCTCCAACCATCATCTGGAA. Primer efficiencies for Gapdh and Ncam1 were determined to be 93.93% and 93.17%, respectively.

3.5. NanoString nCounter Processing

A custom codeset of 35 genes was designed, targeting a region within the amplicon generated from the forward and reverse primers (Table 2.2). An error was discovered in the Pcdh18 codeset sequence and was therefore removed from all downstream analyses. The cDNA of each single cell were analyzed with our custom codeset NanoString Headquarters (Seattle, WA) as per standard NanoString nCounter protocol.

Gene Name	PCR Forward	PCR Reverse	Hybridization Target
Cdh1	GCCCCCATCTTTATGCCT	TGGCACCAGTCTCTGGGT	TCGAAGTGCCCGAAGACTTTGGTGTGGGTCAGGAAATCACATCTTATACCGCTCGAGAGCCGGACACGTTCATGGATCAGAAGATCACGTATCGGATTTG
Cdh13	CGGACCCTGGACAGAGAA	TAGCCGGGTCATCTGCAT	GGCCAGTGCCTCTGGAAGTCATTGTGATTGACCAGAACGACAACAGACCCATCTTCCGGGAAGGCCCTTACATCGGCCATGTCATGGAAGGGTCACCCAC
Cdh2	CGCCATCATCGCTATCCT	TCTGGTTGCTGGAGCTGG	TGATCCTTGTTCTCATGTTTGTGGTATGGATGAAACGGCGGGATAAAGAGCGCCAAGCCAAGCAGCTTTTAATTGACCCAGAAGATGATGTAAGAGATAA
Cnga2	GGACAGGGGCACGATGTA	TGAGCGCTGAAACCTGTG	CTTAGCTGTGGTTATGTGCTCCATTGGTCCTCTGTTGGCTTCCTGGGCTGGTATCAGACAACGTAGGGTAAATCTCTGATTGGGAGGCTGCTGCTGTTTG
Efna5	ACAACGGAAGAAGGTCCTG	ACCGCGGGATGGCTC	TGAGACCAACAAATAGCTGTATGAAAACTATAGGTGTTCATGATCGTGTTTTCGATGTTAACGACAAAGTAGAAAATTCATTAGAACCAGCAGATGACAC
eGFP	TCGTGACCACCCTGACCT	GTTGCCGTCCTCCTTGAA	GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAAC
Epha5	AGGGCCTGAAACCTGCAT	CACTGAGGAGGAAGCCGA	CAAATTCGAGCACGTACAGCAGCAGGCTACGGCGTCTTCAGTCGAAGATTTGAGTTTGAAACCACACCAGTGTCAGTTGCAGCATCTAATGATCAAAGCC
Gap43	AGGAAGAAGGCAGGGGAA	GCTGGTGCATCACCCTTC	TACCACCATGCTGTGCTGTATGAGAAGAACCAAACAGGTTGAAAAGAATGATGAGGACCAAAAGATTGAACAAGATGGTGTCAAGCCGGAAGATAAGGCT
Gapdh	TCTGAGGGCCCACTGAAG	GCCATGTAGGCCATGAGG	AGGTTGTCTCCTGCGACTTCAACAGCAACTCCCACTCTTCCACCTTCGATGCCGGGGCTGGCATTGCTCTCAATGACAACTTTGTCAAGCTCATTTCCTG
Gfap	GGCTGGAGGCAGAGAACA	CGCTGTGAGGTCTGGCTT	CACCCTGGCTCGTGTGGATTTGGAGAGAAAGGTTGAATCGCTGGAGGAGGAGATCCAGTTCTTAAGGAAGATCTATGAGGAGGAAGTTCGAGAACTCCGG
Gnal	GGACACCTAGACACAGCTCACTC	TGAGCATCTCAGCCAGGA	AGTTCCGAGGCTTTGGGGGGTCATTCTTAGGCTTAAGTTATCTGTAAAGCTGTCCAAAGTTAAGTTTTTAACAGTACAAATAGGGGGGAGTGACAGCCTG
Ncam1	AAGAATGCACCAACCCCA	TTGATGCCCCTGATCTGC	ATGCTGTGATTGTCTGTGATGTGGTCAGCTCCCTGCCTCCAACCATCATCTGGAAACACAAAGGCCGAGATGTCATTCTGAAAAAAGACGTCCGGTTCAT
Ncam2	AATCCCATGGAGGTGTGC	ACTTGGCTCACGGACTGG	TGGAAAATAGTACGCTCCCATGGAGTTCAAACGATGGTTGTTTTGAGCAGTCTGGAACCAAATACGACTTACGAAATTAGGGTTGCAGCAGTGAATGGCA
Notch2	TTCTGGAGATCGACAACCG	CCCCAGCAGGATGAAGAA	GGCTTCTCATGCCATCCAAGGGACCCTATCCTACCCTCTAGTGTCTGTTTTCAGTGAACTGGAGAGTCCAAGAAACGCCCAGCTTCTCTATCTGCTCGCG
Nrp1	TGTGGTGGATTCTTCTCGG	CTGCAGGTAGCGATGCAA	CTAGGCGTTATTCAGTGGGCCCTTTTGATGGACATGTGACCTGTAGCCCAGTGCCCAGAGCATATTATCATAACCACATTTCAGGGGACGCCAACGTCCA
Nrp2	CTTTCGGGAAACCCTCGT	TGGGTCCTGGCAGTCAGT	ACATACACGCGTTTGGGTTTGAAGAGGAAACCGGTCTCCGCTTCCTTAGCTTGCTCCCTCTTTGCTGATTTCAAGAGCTATCTCCTAGTGAGGTGGAGAT
Omp	GGTCTGGGCCTGGTTTTT	CCACCAAGTCCACAAGCC	CTCATATGGGTATTGAGAGATGGCATAACTATTTGTATGGATGTGGGCCTGAGGGCTAGTCCTGGTGAGGAGTAAGGCTAACTTTAGTTTAATTATTGAG
Pcdh1	TGACGGAGACAAAGGGGA	GTGGGACTCCGCCATCTA	CTCTCGGTAGAGCAGGACAATGGTGACTTTGTTATCCAGAATGGCACAGGCACCATCTTATCCAGTTTGAGCTTCGATCGGGAGCAACAGAGTACCTACA
Pcdh10	CTGGCTACAGGGGAGCTG	CCACCGTCTGAAACCTGC	TGGAAATGATTGTGCTGTTATTCTTTGCCTTGCTCTGGATGGTGGACGGAGTCTTTTCCCAGCTTCATTATACTGTGCAGGAAGAGCAGGAACATGGCAC
Pcdh11x	GTGAGCCAGAGCAGCCAG	CCCGCTCCACAGAGACTG	CCCAGCTAGCCCGGGCCGCCCGACCAGCCTCACAAACAGCCTCCAAGACAAACAGCTGCTCTGCGGACTGCGAGCTGTGGCGGTAGATCCCGCTACAGCA
Pcdh17	CTGACAGGCACCGCACTA	CAGTCTTAGGGGTGGGGG	CACAATTAGCTATGCACACCGATGCTGGACTTTGTGAACAGAGACCGTTGGGAGGAAAATCTTGAACTATGATCCTGGGCTAAGCTGTGAATCATTACTG
Pcdh19	CATTGCGGGCATTCTCTT	CGAGGTACAAGGCGGATG	CTATCCCTGCACACTGCAAAGTCACAGTCAGCGTACTTGATACTAATGACAACCCGCCAATTATCAACCTCCTGTCGGTCAATAGCGAGCTTGTGGAGGT
Pcdh19	GTCACCGGTGCGCTAGAC	CGATCAGACACCCGAACC	CTGCATTGCATCTCGGTTTCTCCCAATAGTGAGGAGCAAGACAAAAAGGCAGAGGAGAAAGTGAGCCTAAGGGGAAAGAGAATCGCTGAGTACTCCTATG
Pcdh19	TCACTTGTCTCCTCGGCTG	CGAGGTACAAGGCGGATG	GGCAATCAAATGCAAGCGTGACAACAAAGAGATCCGGACCTACAATTGCAGAATCGCTGAGTACTCCTATGGGCATCAAAAGAAATCAAGTAAGAAGAAA
Pcdh7	TCCTACACGTTGCTGCCA	CACAACCAGCCTGTGCAA	ATGCGGACTTGAACTACAGCATTGTGGGAGGGAATCCTTTCAAGCTGTTTGAGATTGATTCCACCAGTGGTGTGGTTTCCTTAGTGGGGAAACTCACCCA
Pcdh8	TTACGGTGCCTCTCCAGG	CAGCAGTACACGCCCACA	TCTGGAAGGGTCATTCATTCAACACCATCTCGGGCCGAGAAGCTGAGAAGTTCAGTGGGAAAGACAGCGGCAAAGGAGACAGTGATTTCAATGACAGTGA
Pcdh9	GGCCGCAATTGATTGTTC	AGAGGTCCCCACAGGAGC	CAGAAATCGAGCACAGCTATTCTCCAGGTCACAGTAAGTGATGTCAATGACAACAGGCCAGTGTTTAAAGAGGGTCAAGTAGAAGTGCATATTCCAGAGA
Rpl13a	ACCAAGAGAGGCCAGGCT	TGCCTGGTACTTCCACCC	ATGGGATCCCTCCACCCTATGACAAGAAAAAGCGGATGGTGGTCCCTGCTGCTCTCAAGGTTGTTCGGCTGAAGCCTACCAGAAAGTTTGCTTACCTGGG
Rpl19	GGCCCACAAGCTCTTTCC	AGTCACAGGCTTGCGGAT	GCCTCTAGTGTCCTCCGCTGCGGGAAAAAGAAGGTCTGGTTGGATCCCAATGAGACCAATGAAATCGCCAATGCCAACTCCCGTCAGCAGATCAGGAAGC
Sdha	GCTGTGGTTGTAGGCGCT	AGCCAGTCGGAGCCTTTC	CTTGCGAGCTGCATTTGGCCTTTCTGAGGCAGGGTTTAATACTGCATGCCTTACAAAGCTCTTTCCTACCCGATCACATACTGTTGCAGCACAGGGAGGT
Sema3a	GCGGGATCCTGTGTCATC	GGCAGTGGCGCTTAAGAA	TCTTTCTCTGTTTCCCTTTCATTCTGCTTCCTCGGAGCCGAATGAAGCAGGGAGAGGGAGCAGGATTAGAGTCAGCCACCGGCTATCAGCGGAGCGGAGA
Sema3c	TCATGGTCCTTGTGGTGC	AAAGCCAGTGACAATTAGAAAGC	CTTCAAAATCAATGCGACAACAGAAGCAGAACTGTTACAGCCTCGGTTGAGCGAGGGCCATAAATTTCCCTGTGCTCTTCCTTCCGTGCTCTAGGGGTTT
Sema3f	TAGGACCAGCCTAGCCCC	GCCCTGGCAACTCCCTA	GCAAAGACAGTATTTATTGGTGGGTTGTATATAGCCTGCCTAGGTGGCAGCATTATCCAGAACTTAGACCCGTGTTGGTCAGCAATGGCACATGACTGAG
Sema4d	CAGTCCCTTGAGGACGGA	TCGCGGGATCATCAACTT	CGTGGCTGAACGAGCCTAGCTTCGTCTTTGCTGACGTGATCCAGAAAAGCCCAGATGGTCCGGAGGGTGAAGATGACAAGGTCTACTTCTTTTTTACGGA

Table 2.1 Custom NanoString codeset sequences.

3.6. Analysis of Single Neuron NanoString Data

Statistical analyses of single-cell data was performed by our collaborators, Dr. Jean Yang and Dr. Shila Ghazanfar in the School of Mathematics and Statistics at the University of Sydney, Australia. We utilized a statistical model that we have previously published (Ghazanfar et al. 2016). Briefly, the distribution of the NanoString $\log_2$ counts for each δ-protocadherin was fitted to a gamma-normal mixture model using an Expectation Maximization algorithm. A maximum-likelihood classifier was then adopted to classify cells as "expressed" or "not expressed", by assigning cells as "expressed" if the fitted normal density for the second component is higher at that expression level than the fitted gamma density for the first component. To ensure that the assigned labels of "expressed" or "not expressed" corresponded to the relevant biological scenario, the mean of the second fitted component was compared to that of negative control gene eGFP via one-sided t-tests. In particular, it is desired that the mean of the second component is significantly greater than the mean of the eGFP expression (calculated as 3.9 - the largest among other negative control genes Notch2, GFAP and Cdh13). As a result, seven protocadherins were studied further, as Pcdh11x and Pcdh18 did not pass this criteria at the significance threshold of $P < 0.05$.

3.7. Single-cell qPCR Validation

Single-cell qPCR validation of the NanoString data was performed identically to the single-cell pre-NanoString quality control. Briefly, 2uL of the cDNA from each single cell were used as template in each Taqman assay (Gapdh, Ncam1, Notch2, and all nine δ-

protocadherins). All primer sets (Table 2.2) displayed efficiencies between 93-100%, except for Pcdh1 which had 83% efficiency (improvement was not observed with multiple primer designs).

Gene	Forward Nested qPCR Primer	Reverse Nested qPCR Primer	Taqman Probe	Efficiency (%)
Gapdh	AACAGCAACTCCCACTCTTC	TTGTCATTGAGAGCAATGCCAGCC	CCTGTTGCTGTAGCCGTATT	93.93
Ncam1	GATGCTGTGATTGTCTGTGAT	GACTATGAACCGGACGTCTTT	CCTGCCTCCAACCATCATCTGGAA	93.17
Pcdh1	CTGTCTGAGAACAGTCCCATTG	AACGACTCAGACCAAGGTGCCAAT	GGTGGAACGTGTAGTCAATCTC	82.74
Pcdh7	CTGGAGACATCCTGGTCAATAC	CCAAAGACAAAGGCATCCCAGTGC	TCAGCCACCTGTACAATCAC	97.49
Pcdh8	GGAGATCATGCTTACTGGAGAC	TGTTCAAGGCCCTGCTGGTCATAT	GTTGCAGTGGTGGTGAGA	100.87
Pcdh9	CCATTCCTGGTTCTGTGGTT	TGCGGTGGATATTGACACCGGAAT	TGTTGTTCCCGCTCACTATG	93.16
Pcdh10	CCCAGTGTATCAGGTGTATGTC	CTGCAAGGTTCTGGTGAGAGTGCT	ACTGTGCTGAAGCTGATCTC	100.04
Pcdh11x	TTGACCCTCCTTCCAATTACTC	TTCTACCATCCACAAACCCTGGCA	CCTCTGCATTCATGCCAATATC	100.4
Pcdh17	CAAGTTTCAGCAGCCAACATAC	TGCAAATTCCCGAGAACAGTCCCA	CCATAATCTCCACTGTCCAAGTC	100.01
Pcdh18	ACCTACTCGCTCTCTGCTAAT	ATCGAAGTGCGAACCAGGACAGAC	CTCCCGATCCAACTCTTTCAC	96.27
Pcdh19	AGCTGTTTGGCCTGGAAATA	TTGCTGAACTGGTGGTGGAGAAGA	CCGTAATGCGAAAGCTGTAATG	96.73

Gene	Forward Flanking RT/PCR Primer	Reverse Flanking RT/PCR Primer
Gapdh	TCTGAGGGCCCACTGAAG	GCCATGTAGGCCATGAGG
Ncam1	AAGAATGCACCAACCCCA	TTGATGCCCCTGATCTGC
Notch2	GATCAATTGCCAAGCAGATGTC	GTCAGTGATGTCCCGGTTG
Pcdh1	GTGCTTGATACAAACGACAACG	ACGGTGATAAGTCCAGTGTTTC
Pcdh7	ACCTACTCGCTCTCTGCTAAT	TCTTGCATAAACTTAGGGTCGT
Pcdh8	GGGAACTGGCCTTTGATCTG	CCCGGCTGTTACCACGAA
Pcdh9	AGTTGTCATTTCTCCACCATCC	TAGGTGCTGGCTTCTCTTCTA
Pcdh10	GCGAGCTGGACTATGAAGAG	GATCGGTCACGCTGAACA
Pcdh11x	GTATCACGTTCTTCAACTGCTAGA	GTATTTCCTCCAACAATGCTGTAAA
Pcdh17	GACGGCACAGGTTCTCATT	CCCGGTTACAGGGTTTACTTC
Pcdh18	GCGAGCTGGACTATGAAGAG	GGTGAGCTGAAGCTCGTAG
Pcdh19	TTGGTGTGCAGACCTACGA	GGTTGTTGTCATTCGAATCGGT

Table 2.2 Primer sequences for qPCR screen and validation experiments.

4. Results

4.1. δ-Protocadherin Expression in the Mouse Olfactory Epithelium

We first performed single color digoxigenin-labeled *in situ* RNA hybridizations for each δ-protocadherin in the mouse olfactory epithelium, as well as neural (Ncam1) and glial (Notch2) markers (Figure 2.1). Ncam1 signal is consistently strong throughout a layer in the epithelium, marking the neural layer, while Notch2 marks the non-neural sustentacular cells (Rodriguez et al. 2008). Of the nine δ-protocadherins, we detected seven in the neural layer of the olfactory epithelium (Pcdh11x and Pcdh18 were not detected). The individual pattern of each delta varied. For example, Pcdh7 and Pcdh17 were both strongly expressed overall, but in some regions are weakly expressed. In contrast, Pcdh8 appears to be almost uniformly expressed throughout the entire epithelium, but only in the most apical region of the neural layer and therefore appears to have weaker expression overall.

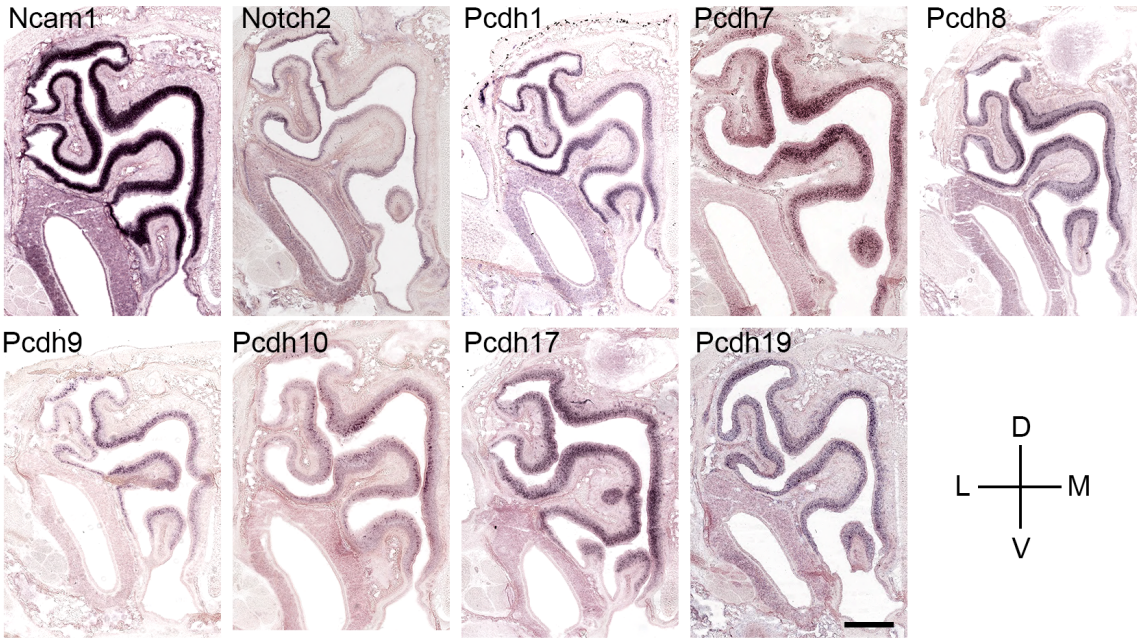

Figure 2.1 RNA *in situ* hybridization of the mouse olfactory epithelium at P7 for Ncam1, Notch2, and seven of the nine δ-protocadherins. The remaining two, Pcdh11x and Pcdh18, were not detected. Note the general expression for Ncam1 and Notch2, whereas the δ-protocadherins exhibit restricted and punctate patterns. Each *in situ* was performed over three different animals, and control sense probes did not generate any signal. Scale bar = 500 µm.

Interestingly, some deltas had notable patterns of spatial restriction. For example, Pcdh1 was consistently expressed the highest in the dorsal-lateral region, with several other sub-regions demonstrating localized variable expression patterns (Figure 2.2A, see arrow heads). In contrast, Pcdh7 had minimal expression in the dorsal-lateral region, and had strongest expression in between the later and medial regions. Finally, within all areas of expression, signal always appeared punctate (Figure 2.2B), suggesting that only a proportion of the cells in a region being expressed are actually expressed. This is consistent with double label *in situ* hybridization data from our lab

(unpublished, Eric Williams) on E17 mouse olfactory epithelium, which shows that these seven deltas can all co-express with each other to various degrees. Together, these *in situ* studies are concordant with expression studies on δ-protocadherins in other regions of the nervous system (Luckner et al. 2001; Krishna-K et al. 2008; Hertel et al. 2008; Krishna-K et al. 2011; Lin et al. 2012; Hertel et al. 2012), and also with the hypothesis of delta combinatorial expression in single neurons. However, these methods are unable to define combinations and reveal how extensive they may be.

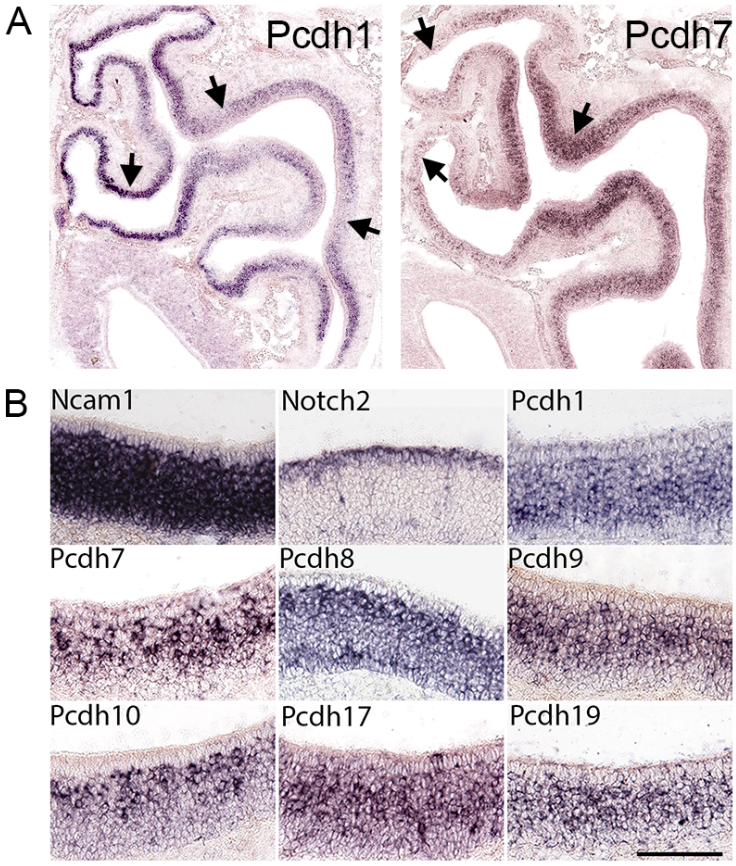

Figure 2.2 RNA *in situ* hybridizations of the mouse olfactory epithelium at P7 from Figure 2.1, but highlighting spatial regulation and punctate patterns. (**A**) Some of the deltas exhibit strong spatial regulation, as seen here by Pcdh1 and Pcdh7 (see arrow heads). (**B**) Higher magnification images at the midline showing punctate expression of the deltas, but not for Ncam1 and Notch2. Scale bar = 100 µm.

4.2. Development and Optimization of Single OSN RNA Processing for NanoString Detection Analysis

Single-cell analysis using NanoString for detection of gene expression has not been established as a routine method. Therefore, careful optimization at each step was required to develop a robust and reliable method.

4.2.1. Strategy and Optimization of Single OSN Processing

The overall strategy for processing single cells was to maximize sensitivity by only targeting specific genes, while minimizing noise from amplification biases, contamination and excess handling. Furthermore, we predicted that by only targeting small regions of specific genes of interest and using a non-sequencing based method of detection, we may reduce amplification biases while increasing sensitivity of our genes of interest. To this end, we introduced a multiplex pool of forward and reverse for primers for all 35 genes at the first step of reverse transcription. This way PCR amplifications could commence immediately after first strand synthesis without further manipulations. The technical aspects of single OSN processing required exhaustive optimization to produce strong and reliable signals. For optimization experiments, dilutions of bulk RNA at various amounts (e.g. 1, 0.5, and 0.1 pg of total RNA) as well as single OSNs were used. To assess the outcome of optimization experiments, the cDNA produced was assayed using nested Taqman qPCR for the marker genes Gapdh (indicating a cell was captured) and Ncam1 (confirming it was a neuron). The following conditions were exhaustively optimized for the processing of single cell mRNA: method

of single cell isolation, type of lysis buffer, volume of lysis buffer, multiplex primer concentration for reverse transcription, denaturing temperature of RNA prior to reverse transcription, reverse transcription buffer, type of reverse transcriptase, units of reverse transcriptase, reverse transcription temperature, reverse transcription time, cleaning up reverse transcription products prior to amplification, number of PCR cycles, units of PaqTaq in PCR amplification, Mg^{2+} concentration in the PCR reaction, as well as others. These parameters were carefully optimized to prevent exponential distortion of amplicons. The final processing method is outlined in the methods section.

4.2.2. Optimization of Single-Cell cDNA Input for NanoString nCounter Detection

Once the single-cell processing protocol was finalized, total signal from any cell could be increased or decreased by changing the number of cycles in the amplification step. It was important for each cell's cDNA to be within NanoString nCounter's dynamic range of detection such that no cell would saturate or fall below detection of the system. The optimal signal was empirically determined by running cells with different qPCR signals (from endogenous variation and from different amplification cycles) for the marker genes on NanoString.

As a general rule, we found that any cell with a Ct value above 25 would not have strong NanoString detection and that lowly expressed genes would risk being excluded from detection. Saturation was not approached until the signal became much stronger (below Ct values ~16). Therefore, as a measure of quality control, the cDNA of each cell was screened for Gapdh and Ncam1 prior to analysis with NanoString. Only cells with

Gapdh and Ncam1 qPCR Ct values of 25 or lower were sent for further analysis with NanoString.

Out of 155 OSNs isolated, processed and screened, only 61 passed our quality control measure, indicating a success rate of 39% (Figure 2.3). Although each cell was manually isolated under a microscope, it is still possible that cellular content was lost in the steps of micro-manipulation. PCR failure did not seem like a probable factor and was ruled out because qPCR replicates were reliably consistent and never failed over multiple re-runs. We therefore believe that the failure of most cells occurs in the process of isolating and capturing the cells.

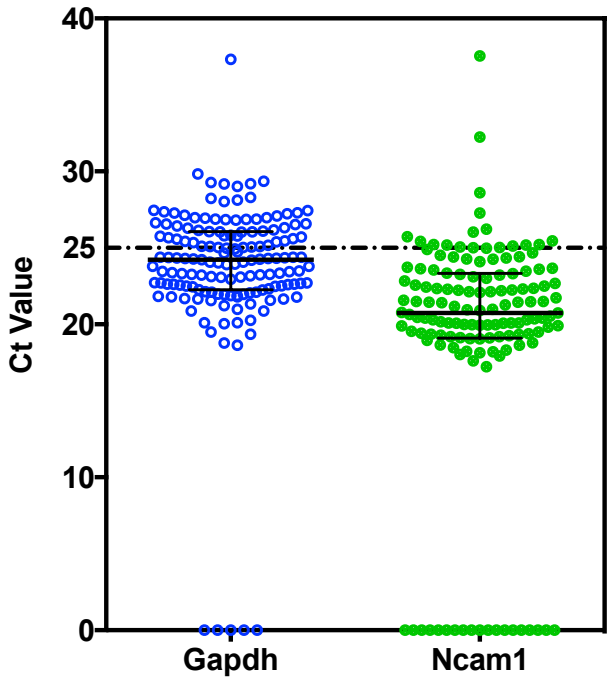

Figure 2.3 Distribution of qPCR quality control screen for 155 single OSNs for Gapdh and Ncam1. Cells with Ct values of 0 (x-axis) did not have any detectable signal. Only cells with Ct values ≤ 25 (dotted line) for both genes were used in further analysis with NanoString.

In total, 50 single cells as well as several controls were analyzed on the NanoString nCounter for our codeset. Controls included two water only samples, and four bulk samples of 100ng of purified RNA from 1,000s of OSNs. In single-cell analyses, it is not possible to perform biological replicates, because the fundamental unit being measured is the cell itself. Therefore, in order to estimate technical noise and replicability, we also ran four "pool/split" samples. These were prepared by combining and mixing 12 individually isolated and lysed cells, and then distributing the pooled volume back into separate tubes of equal volumes and proceeding with processing them as if each was a single cell. If technical replicability of the single-cell processing and NanoString is high, these samples should produce very similar results.

To validate the use of our quality control screen, we first compared each cell that was run on NanoString to its corresponding qPCR values for Ncam1 and Gapdh (Figure 2.4). Log_2 counts from NanoString showed high correlation to $1/log_2$ of the Ct values, with Pearson correlation values of 0.89 (P-value $< 1\times10^{-15}$) for Ncam1 and 0.86 (P-value $< 1\times10^{-15}$) for Gapdh. This high concordance suggested that the digital count output from NanoString is as sensitive as qPCR and does not require further standardization or normalization and also that the relationship between them appears to be linear. Therefore, further analysis on the NanoString dataset was performed on the log_2 counts.

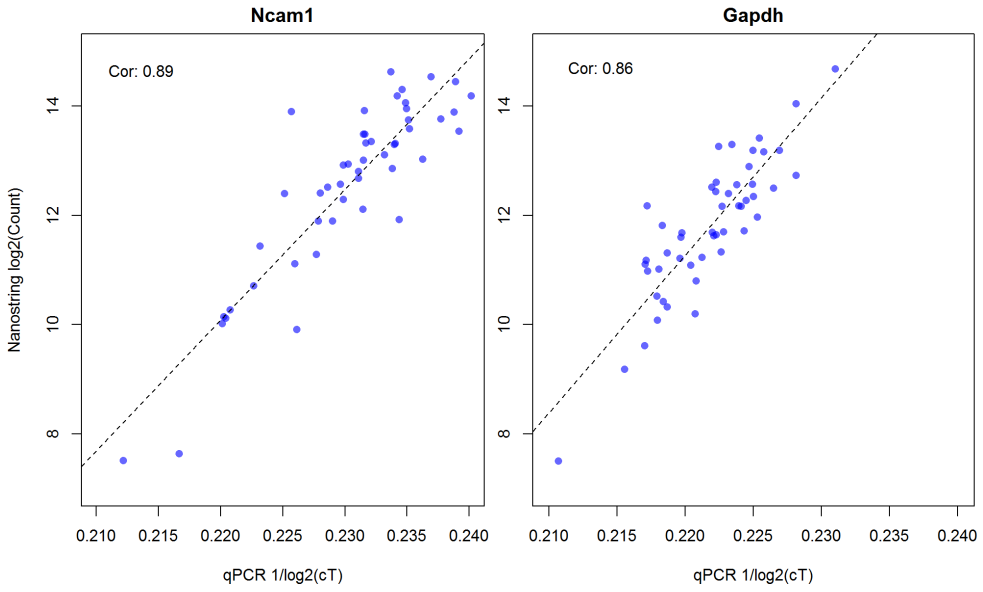

Figure 2.4 Pearson's correlation coefficient of single OSN qPCR Ct values to NanoString counts for Ncam1 (left) and Gapdh (right). Both display high concordance, indicating that the digital count output from NanoString is as sensitive as qPCR and no normalization is needed.

4.3. Single-Cell Analysis of the δ-Protocadherins

4.3.1. Single-Cell Heterogeneity and Sample Controls

Strong heterogeneity among the single cells can be seen in the heat map of raw NanoString counts (Figure 2.5). This is consistent with the proposed combinatorial expression of δ-protocadherins as well as other guidance cues. As expected, the water only samples were negative, confirming the lack of contamination or background signal. Comparing the average of single cells to bulk RNA is a common way to gauge if sufficient single cells have been sampled and if they accurately reflect the tissue's composite gene expression profile. Although our bulk samples are comprised primarily of neurons, there are also some non-neuronal cells within the populations, causing a

few discrepancies in signal comparison. Additionally, our single OSNs were selected for strong Ncam1 signal, whereas the neurons in the bulk samples were not.

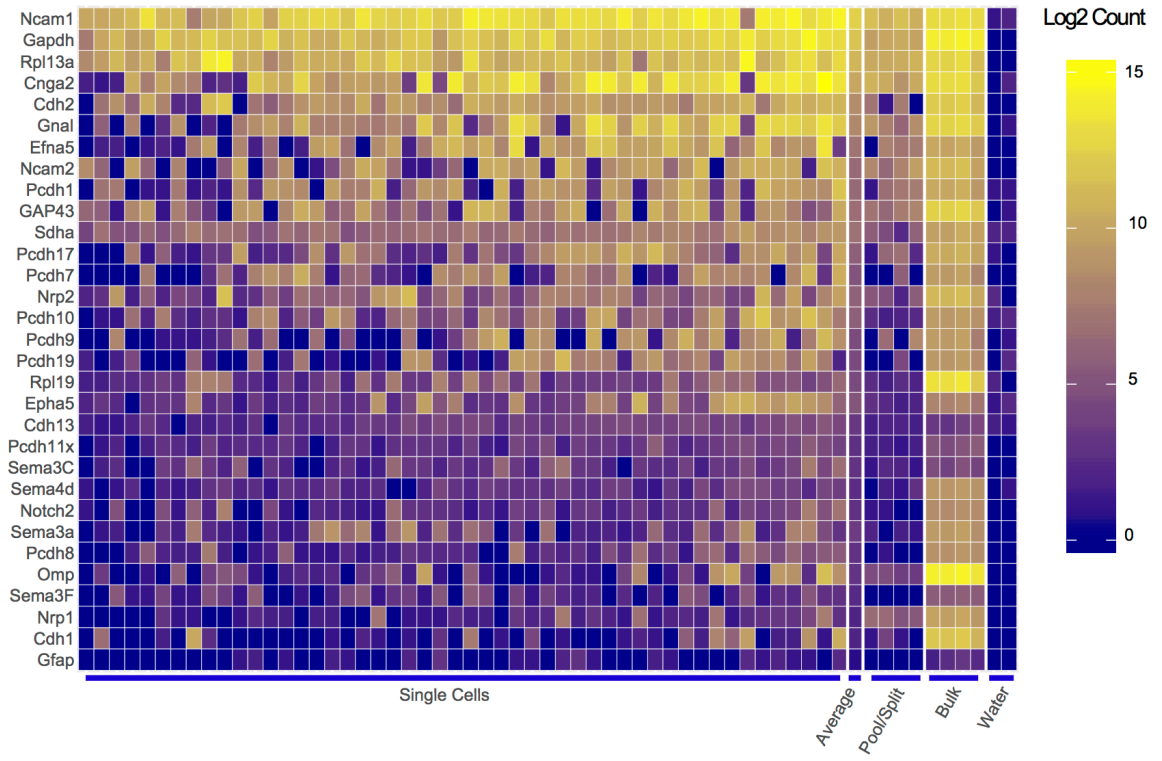

Figure 2.5 Heat map of raw NanoString $\log_2$ counts shows heterogeneity of deltas and other axon guidance cues in codeset in single OSNs.

Using Multiple Discriminant Analysis (performed by our collaborators Jean Yang and Shila Ghazanfar at the University of Sydney, Australia), we examined whether the four sample types (single cells, pool/split, water, and bulk) were more similar to themselves then samples in other groups (Figure 2.6). Across the first two dimensions, each sample type occupied a distinct dimensional space. The single cells occupied the most disperse space, indicating high levels of heterogeneity and the possibility of sub-populations within the 50 cells. The pool/split samples fell within the single-cell space, but clustered

together tightly in one region, indicating high technical replicability. Finally, the water and bulk samples were well separated from the single-cell and pool/split samples.

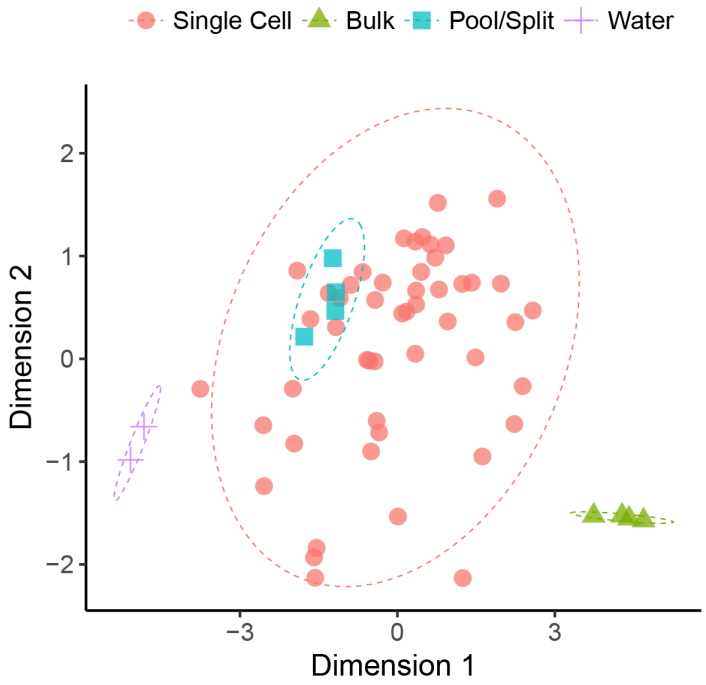

Figure 2.6 Multiple Discriminant Analysis showing clustering of sample types.

Since single cells must be amplified, a major concern is the possibility of generating erroneous signal through amplification biases or exponential distortions. The ribbon plot in Figure 2.7 shows that despite our bulk sample containing non-neuronal cells, that the single-cell populations demonstrate similar trends with the bulk. Notable differences between OSNs and the bulk are observed in olfactory marker protein (OMP), E-cadherin (Cdh1), and Neuropilin1 (Nrp). These differences are likely caused by sample size, the fact that non-neural cells are present in the bulk, and that the single cells were screened and selected for high Ncam1 signal. Despite these differences, strong

correlations were found between sample types, and the single cells showed strong correlation to the bulk samples ($r^2 = 0.65$), and high correlation with the pool/split samples ($r^2 = 0.87$).

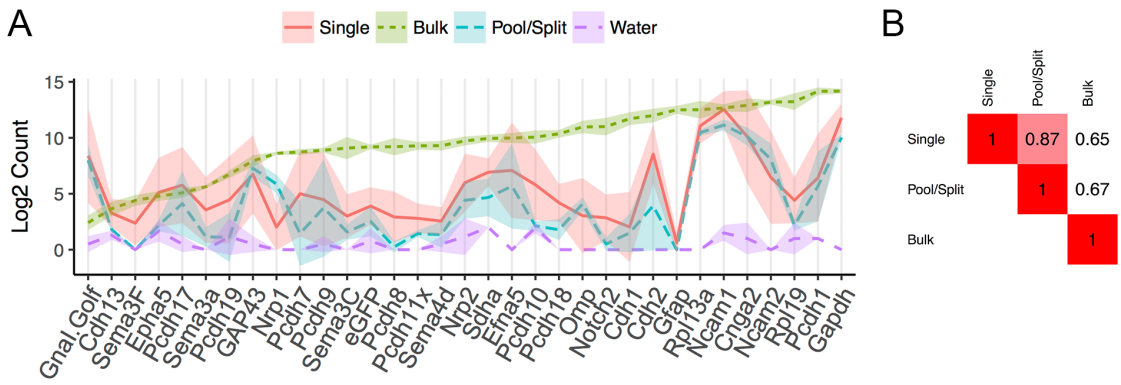

Figure 2.7 Comparison of sample type by expression rankings. (**A**) Ribbon plot showing average expression for each gene from the single cells, bulk, pool/split and water samples. The overall trends track well except for a few outliers that are not well detected in the single cells. (**B**) Pearson's correlation between sample types. The single cells display high correlation to the bulk, indicating that the single cells appropriately represent the bulk tissue, and the pool/split show strong correlation to the single cells, indicating high technical reproducibility.

4.3.2. Mixture-Modeling to Determine On/Off Status of δ-Protocadherins

The relatively limited number of neurons tested prevented us from accurately assessing quantified variation of deltas between the cells. Instead, to define which deltas were expressed in any given OSNs with high statistical confidence, we utilized a gamma mixture model approach that we previously published (Ghazanfar et al. 2016) to classify each delta as "on" or "off" in each cell (see methods, performed by our collaborators Jean Yang and Shila Ghazanfar at the University of Sydney, Australia). This analysis

showed that individual OSNs can express anywhere from zero to seven deltas. This is consistent with the presence of a combinatorial code for δ-protocadherins, as proposed elsewhere in the CNS. Although Pcdh18 was excluded from the codeset, Pcdh11x was not detectable in any single OSNs, consistent with our *in situ* hybridizations.

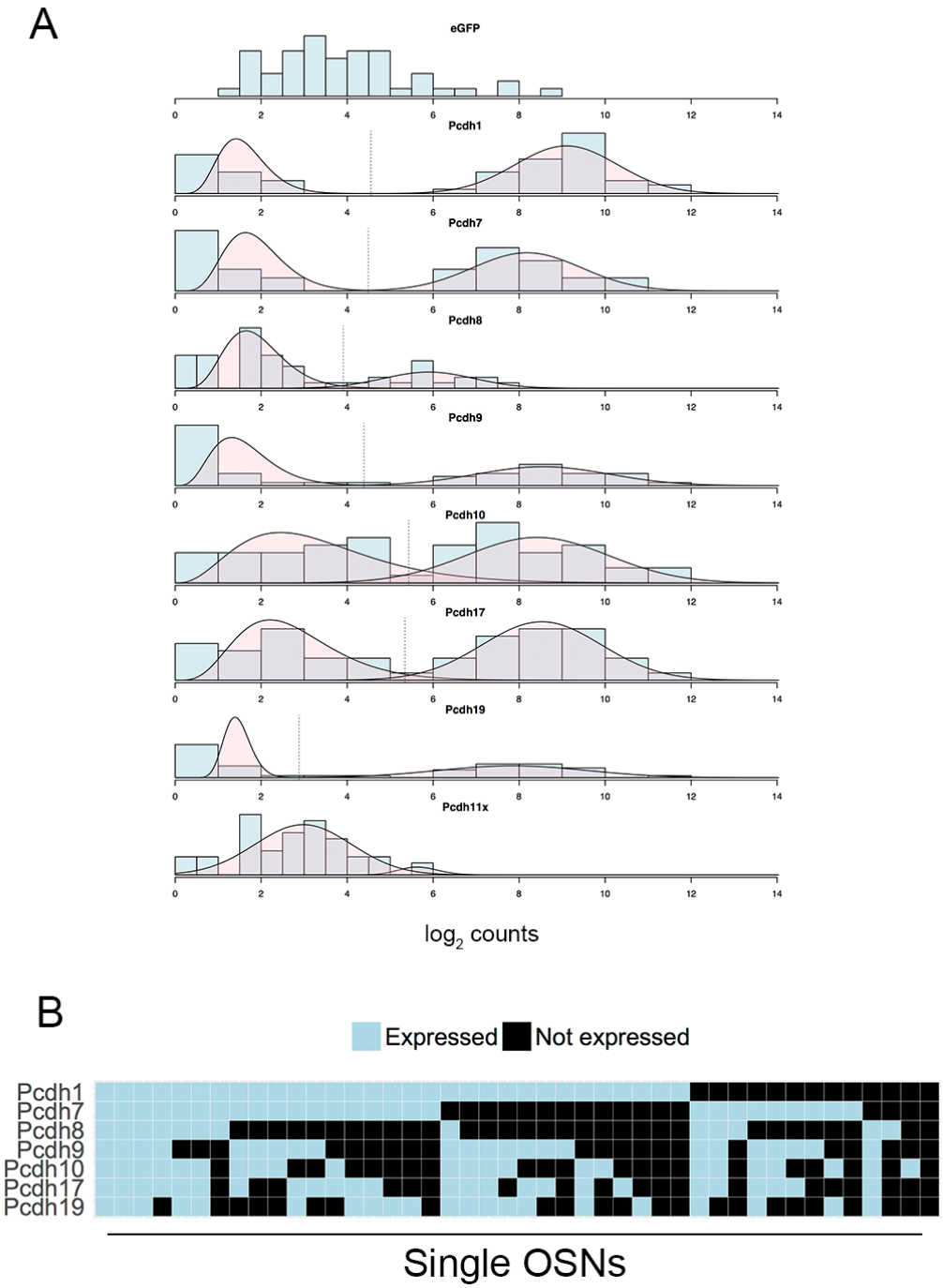

Figure 2.8 Gamma mixture-modeling for calling on/off status for each δ-protocadherin. (**A**) Fitting of mixture-model to each δ-protocadherin. (**B**) Binary heat-map showing the mixture modeling results for the δ-protocadherins, indicating a wide range of combinatorial states can be expressed in single OSNs.

4.4. Validation of Single-Cell Data

NanoString is a relatively new platform for single-cell analysis. We therefore sought to validate our results by utilizing single-cell qRT-PCR on a population of 18 randomly isolated individual OSNs (Figure 2.9). We designed Taqman primers that targeted genomic regions that were independent from the NanoString codeset targets. We utilized a similar “on” or “off” call for the various deltas by establishing a threshold Ct value of 30, where any gene with Ct > 30 was called “off” and any gene with a Ct ≤ 30 was “on". This analysis showed that the pattern of expression shown by qRT-PCR is consistent with the combinatorial patterns in NanoString data, and that no cells expressed Pcdh18 and only a small percentage expressed Pcdh11x.

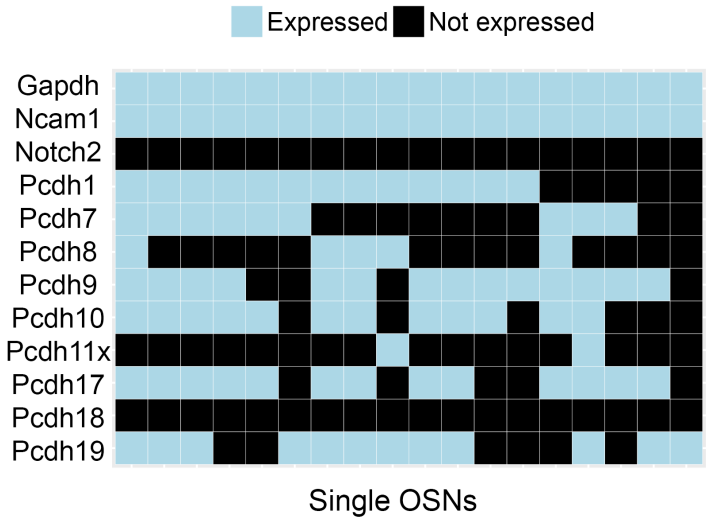

Figure 2.9 Single-cell qPCR screen of the for each delta in 18 single OSNs. Note the similar expression patterns to the NanoString on/off expression heat map and the range of possible combinations.

Finally, a major concern in single-cell studies is whether or not gene expression changes while the cells are being isolated. Before single OSNs can be isolated, the epithelium must be first be dissected and dissociated. To test whether or our single-cell expression data has *in vivo* relevance, we analyzed the *in situ* data from Figure 2.1 to estimate their relative expression levels. The olfactory epithelial layer, where the OSNs reside, was manually traced and the digoxigenin/alkaline phosphatase signal was converted into a percent area positive. This would provide an estimate for the percent of OSNs expected to express a given δ-protocadherin *in vivo.* To avoid biases and arbitrariness in assigning thresholds, each trace was measured at two reasonable thresholds (high and low) to produce a range to approximate δ-protocadherin signal. We then produced confidence intervals from our single-cell data to estimate the likelihood that a random cell would express each delta. Although these are two entirely different methods, they produced similar trends in relative expression values (Figure 2.10). This suggests that the gene expression of deltas did not significantly change while the tissue was being processed or while the cells were being isolated.

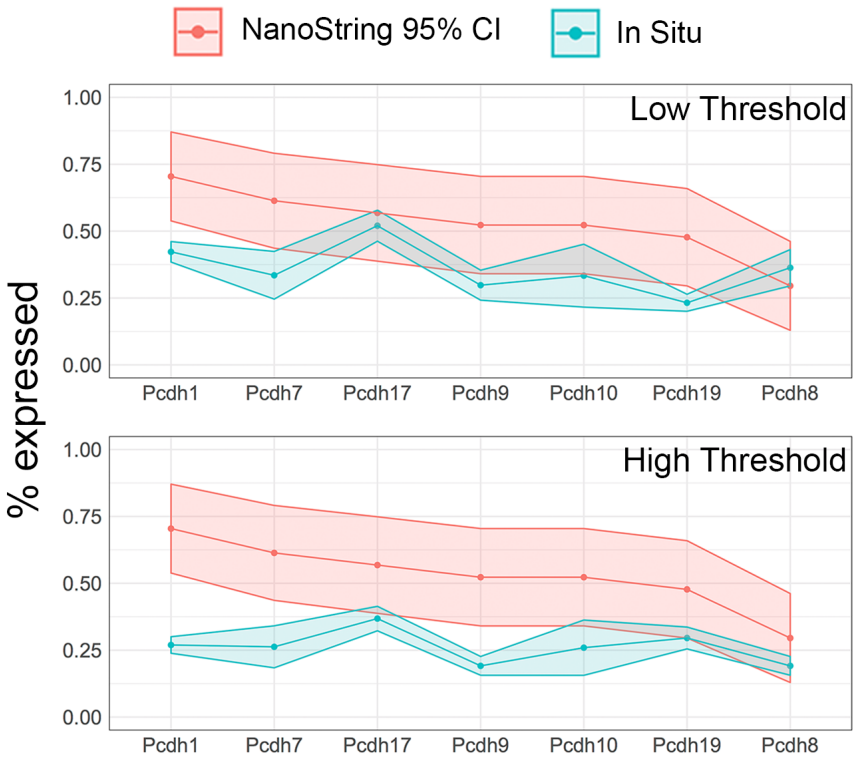

Figure 2.10 Ribbon Plot comparing quantification of *in situ* hybridization signal to 95% confidence intervals generated from the NanoString single OSN results. The red line and area represents the percent of single cells that would be predicted to express each delta in a given set of neurons. The light blue lines represent the mean (and standard error) percent of OSNs that express each delta in the olfactory epithelium *in vivo,* as determined from quantification of the *in situ* signal (n = 3 animals). Because determination of thresholding can easily lead to biases, two measurements were taken for each sample (low and high).

4.5. Comparing the Detection of δ-Protocadherins in scRNA-Seq to Our NanoString Method

Given that scRNA-seq has known limitations in capture efficiencies, we hypothesized that our targeted approach which only reverse transcribes and amplifies the 35 genes in our codeset, would better detect the δ-protocadherins. We sought to estimate if our

targeted single-cell approach using NanoString was better at detecting the δ-protocadherins than scRNA-seq by utilizing three publicly available OSN scRNA-seq datasets (Saraiva et al. 2015; Tan et al. 2015; Hanchate et al. 2015) and analyzed them using the same gamma-normal mixture modeling method (Ghazanfar et al. 2016). Since our single cells analyzed on NanoString were pre-selected for Ncam1 positive neurons, we only analyzed single cells from the scRNA-seq data sets that were positive for Ncam1 to ensure the samples were matched. Comparing the distribution of the mean number of deltas detected per cell (Figure 2.11A) shows that the scRNA-seq datasets often did not detect any deltas, and rarely detected more than two, whereas our NanoString data displayed a wide range in the distribution. The mean number of deltas detected per cell was much lower in all three scRNA-seq data sets than in our NanoString data (Figure 2.11B), with our NanoString detecting an average of 3.66 deltas per OSN while the scRNA-seq data sets ranged from 0.94 to 1.38. Our mean value of 3.66 is more consistent with all of the RNA *in situ* hybridizations experiments done in this chapter, as well as the work done by others in across different neural systems. Together, these are consistent with our targeted single-cell method exhibiting higher sensitivity than existing scRNA-seq methods.

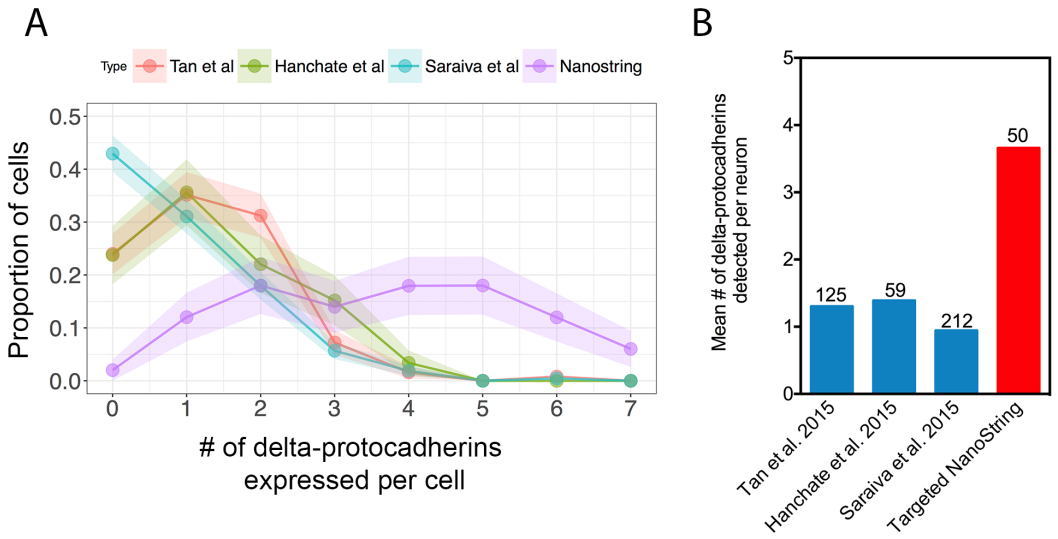

Figure 2.11 Comparison of our single OSN data to three scRNA-seq data sets. (**A**) Distribution of the number of deltas determined to express per neuron. Note that the three scRNA-seq sets are all highly skewed to the left, and rarely detected more than 3 per cell. In contrast, our data detects a much broader distribution. (**B**) The mean number of deltas detected per cell. While the scRNA-seq papers detected ~1 per cell on average, our method detected 3.66. The number above each bar indicates the number of neurons analyzed.

5. Discussion

5.1. Combinatorial Expression of δ-Protocadherins

The cadherins and protocadherins have long been proposed to provide adhesive specificities in the developing nervous system, mediated through combinatorial expression within individual neurons (Shapiro and Colman 1999; Yagi 2012). However, it has only been suggested, and never proven for the δ-protocadherins (Obst-Pernberg and Redies 1999; Hertel et al. 2008; Krishna-K et al. 2011). Defining the combinatorial nature of the deltas is a critical first step towards understanding their functional

implications in neural development because it would guide future studies with a realistic and accurate model for how they are expressed and exist within neurons.

Here, we developed and utilized a targeted medium-throughput, non-sequencing based single-cell analysis approach to show for the first time that δ-protocadherins are expressed in a wide range of combinations in randomly selected olfactory sensory neurons. Although caution must be taken in generalizing findings from one region of the nervous system to others, the overlap and punctate expression pattern of deltas in other regions would suggest that similar levels of combinatorial expression exists in most other neuronal types.

These findings raise several important questions. For example, what is the functional impact on a neuron that expresses multiple deltas? Are some deltas stronger or more important than others and therefore the specific repertoire expressed matters, or are they all equal and the total number expressed is what is important? Addressing these questions will be pivotal in revealing the role of δ-protocadherins in neural development and neural circuit formation.

5.2. Considerations on the Possibilities of Regulated and Stochastic Models of Combinatorial Expression

An obvious question that arises from these combinatorial expression patterns is whether or not they are governed by regulated or stochastic processes. A few plausible models could be imagined. In one case, each delta could be regulated independently, with no regulatory "cross-talk" from each other, and transcription of each delta occurs through

its own independent regulatory mechanisms. In a second scenario, there could be strong co-regulatory aspects that govern which, and how many deltas are expressed in a cell. For example, if certain deltas shared regulatory elements, they may be more frequently co-expressed together than with others. Although we do not have a large enough sample to make conclusions regarding this possibility, we do not see obvious biases of co-expression between different deltas, suggesting that this is not likely. In the third model, delta expression may be entirely stochastic. In a fourth possibility, aspects of these models could be combined. For example, within individual neurons, the total number of deltas to be expressed might be regulated, but their selection is random, or some deltas may be regulated while others are stochastically expressed.

There is another sub-family of protocadherins, the clustered protocadherins, whose regulatory mechanisms have been well studied and may provide mechanistic insight into the possible ways in which the deltas are regulated. For example, each alpha is preceded by its own promoter, and regulatory elements such as transcriptional enhancers that mediate a promoter choice mechanism have been identified and validated (Tasic et al. 2002; Ribich et al. 2006; Kehayova et al. 2011; Monahan et al. 2012; Guo et al. 2012). Regulatory elements have also been discovered for the betas and some of the gammas, but interestingly were not proximal to the genes and could be located over 320kb away (Wang et al. 2002; Yokota et al. 2011). Consistent with a promoter choice mechanism, three separate studies examined the clustered protocadherins using single-cell RT-PCR (Esumi et al. 2005; Kaneko et al. 2006; Hirano et al. 2012). Although not all of the members were tested, they concluded that the

clustered protocadherins exhibit monoallelic yet stochastic expression within individual Purkinje neurons, consistent within the regulatory mechanisms discovered. However, how universal these mechanisms are remains to be seen. For example, from these RT-PCR studies, they found that the C-type isoforms of the alphas and gammas were constitutively expressed in all Purkinje neurons. Despite a lack of evidence, this was perpetuated in the field to be a universal feature of all neurons until recent single-cell studies in olfactory sensory neurons and serotonergic neurons proved this was not true (Chen et al. 2017; Mountoufaris et al. 2017). Thus, discoveries on gene expression and regulatory mechanisms in one neuronal type must be cautiously applied to others.

In contrast, almost nothing is known regarding the regulation of the δ-protocadherins. Given that they are scattered throughout the genome, it becomes more difficult to find and validate any putative regulatory elements across the entire family. To begin approaching this question *in silico*, we examined the distribution of the total number of deltas expressed per cell (performed by our collaborators Jean Yang and Shila Ghazanfar, University of Sydney, Australia). First, an estimate of variation in observed frequencies was obtained by bootstrapping the single-cell data for 50 cells 10,000 times. Next, to model a random distribution based on the observed frequencies for each delta, the average proportion of each protocadherin across the 50 cells were calculated, and then expression status was assigned according to probabilities estimated and also run over 10,000 simulations (Figure 2.12). We found that the middle range of 3 and 4 genes was slightly lower than what would be expected from a random distribution. A

goodness of fit test for these distributions revealed a P-value of $5x10^{-5}$, indicating a significant difference between the two distributions.

This suggests that some level of regulation may be occurring, and that would be consistent with the spatially restricted patterns observed by the *in situ* hybridizations in the epithelium and also elsewhere in the nervous system by others. However, one possible interpretation is that both processes occur. In such a model, deltas could be regulated to only express in certain spatial regions, but within a region that is regulated to be "on", it is stochastically expressed within neurons. Future studies will need to be done to better determine the contributions these processes have on δ-protocadherin expression. How would such an understanding be interpreted towards *in vivo* significance? One obvious method would be to study δ-protocadherin expression across functionally similar neural sub-populations (such as odorant receptor populations) to provide insight into whether specific combinations are expressed by sub-populations of neurons and therefore likely have functional relevance *in vivo*.

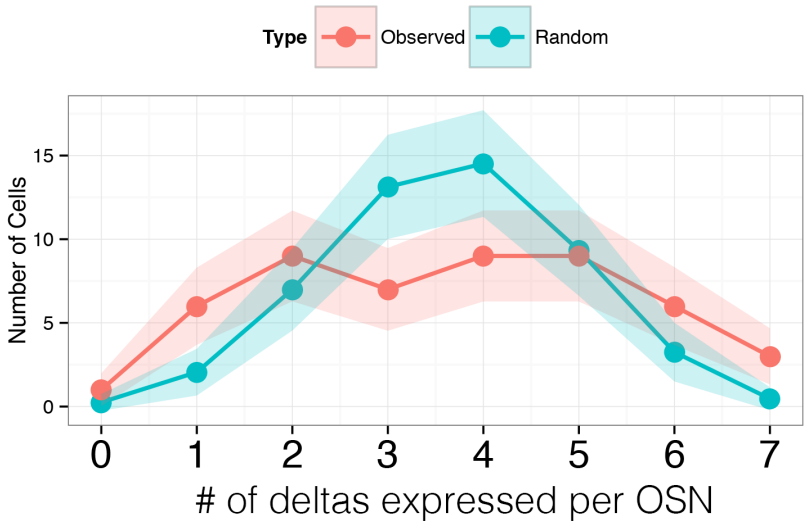

Figure 2.12 Distribution of the mean number of deltas expressed per OSN for observed results and a randomly modeled distribution based on individual expression proportions observed. Note the difference in the middle of the distributions.

To search for any possible conserved regulatory elements among the deltas, MUSCLE sequence alignments of the genomic regions 200bp and 2Kb upstream of each delta transcription start site (TSS) were performed. No sequences or motifs were found to be highly conserved, suggesting that no obvious 5' regulatory elements are shared among the deltas. Interestingly, many lncRNAs were found to be in proximity of the TSS of the deltas (Figure 2.13). In several cases, they are immediately upstream within ~100-300 bp of the delta TSS, but in some cases were slightly further away or intragenic, all consistent with possible modes in which lncRNAs have been found to regulate genes (Kung et al. 2013). While this is striking, it does not appear to be conserved in all mammals. Future studies will be needed to examine the possible role these lncRNAs could have in regulating the δ-protocadherins. Additionally, thorough comparative

sequence analyses will need to be carried out to identified other possible conserved regulatory elements.

Figure 2.13 Examples of lncRNAs found immediately upstream of the transcription start sites of Pcdh1, Pcdh8, and Pcdh10 in the mouse.

5.3. Single-Cell Analyses: Transcriptome vs Targeted Profiling

The field of single-cell biology has expanded profoundly over the past several years, and is becoming more common in biological research as technical and economic barriers are lowered. However, despite its great power to reveal new depths of complexity in biology, it suffers a limitation similar to the human genome project – how can all the sequencing data be interpreted and utilized in a useful and meaningful way? For example, revealing sub-populations of neurons in the brain helps us understand how many different cell types there may be, but it doesn't provide any functional explanation or validation that these cell types are actually functioning differently in the

brain. How can we move beyond clustering and classifications, and towards functional understanding of scRNA-seq data instead?

A well-recognized concern about scRNA-seq is the low capture efficiency of mRNA, and this limitation would theoretically significantly hinder any attempts to interpret the functional implications of a single cell's sequenced transcriptome. While hierarchical clustering may not be greatly affected by this because there are enough genes to still differentiate sub-populations, attempts to functionally understand and validate specific genes may be profoundly impacted by this limitation.

Since we were only interested in a relatively small number of genes, we hypothesized that a targeted gene approach would provide more accurate data by reducing the level of capture inefficiency during the reverse transcription step. Furthermore, the NanoString based approach analyzes small (200 bp) amplicons, and this uniformity would conceivably also help reduced biases among the processing of transcripts. By comparing our data to other scRNA-seq data sets on OSNs, we conclude that our multiplexed NanoString method appears to have a higher detection rate of δ-protocadherins, consistent with our concern about the low capture efficiency of scRNA-seq.

To this end, single-cell experiments should be designed carefully. If the goal is to obtain a general survey of the transcriptomic landscape to identify global changes or sub-populations through hierarchical clustering, scRNA-seq is suitable because it samples the entire transcriptome. However, if you are interested in a limited set of specific

genes, scRNA-seq may not provide enough accurate detection and an alternative method like our targeted approach using NanoString may be more appropriate. Future studies performing direct comparisons will be useful in determining the exact extent of these differences.

5.4. Conclusion

This study developed a novel single-cell RNA analysis method that was used to reveal the complexity of combinatorial expression of the δ-protocadherins. We found that they can be expressed in diverse sets of combinations. This has important implications for understanding their function, as all studies to date have only considered their isolated functions. This knowledge can now be applied to future studies to understand how their adhesive function may be altered when expressed in combinations. Future studies to better understand these combinatorial patterns in the nervous system will require sampling larger numbers of neurons to be able to determine if co-expression regulation is occurring, and testing more types of neurons to determine if these patterns occur throughout all neural cell types. Finally, we compared our results with three publicly available OSN scRNA-seq datasets. We found that on average, our targeted approach detected more deltas per neuron. This raises important questions about the reliability of scRNA-seq results, and whether or not it is suitable for studies that require accurate detection of a smaller number of specific genes.

6. References

Baro DJ, Levini RM, Kim MT, et al (1997) Quantitative single-cell-reverse transcription-PCR demonstrates that A-current magnitude varies as a linear function of shal gene expression in identified stomatogastric neurons. Journal of Neuroscience 17:6597–6610. doi: 10.1016/S0006-3495(93)81431-0

Bononi J, Cole A, Tewson P, et al (2008) Chicken protocadherin-1 functions to localize neural crest cells to the dorsal root ganglia during PNS formation. Mechanisms of Development 125:1033–1047. doi: 10.1016/j.mod.2008.07.007

Bucan M, Abrahams BS, Wang K, et al (2009) Genome-Wide Analyses of Exonic Copy Number Variants in a Family-Based Study Point to Novel Autism Susceptibility Genes. PLOS Genet 5:e1000536. doi: 10.1371/journal.pgen.1000536

Butler M, Rafi S, Hossain W, et al (2015) Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. International journal of molecular sciences 16:1312–1335. doi: 10.3390/ijms16011312

Chen WV, Nwakeze CL, Denny CA, et al (2017) Pcdhαc2 is required for axonal tiling and assembly of serotonergic circuitries in mice. Science 356:406–411. doi: 10.1126/science.aal3231

Chung W, Eum HH, Lee H-O, et al (2017) Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nature Communications 8:ncomms15081. doi: 10.1038/ncomms15081

Citri A, Pang ZP, Sudhof TC, et al (2011) Comprehensive qPCR profiling of gene expression in single neuronal cells. Nature protocols 7:1–10. doi: 10.1038/nprot.2011.430

Crino PB, Trojanowski JQ, Dichter MA, Eberwine J (1996) Embryonic Neuronal Markers in Tuberous Sclerosis: Single-Cell Molecular Pathology. Proceedings of the National Academy of Sciences of the United States of America 93:14152–14157. doi: 10.2307/41042?ref=search-gateway:d4569246e151846a9c370aa66b811f66

Dibbens LM, Tarpey PS, Hynes K, et al (2008) X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. Nature genetics 40:776–781. doi: 10.1038/ng.149

Edström JE (1964) Microextraction and microelectrophoresis for determination and analysis of nucleic acids in isolated cellular units. Methods in Cell Biology. doi: 10.1016/S0091-679X(08)62104-4

Emond MR, Biswas S, Blevins CJ, and Jontes JD (2011). A complex of Protocadherin-19 and N-cadherin mediates a novel mechanism of cell adhesion. J. Cell Biol. 195, 1115–1121.

Emond MR, Biswas S, and Jontes JD (2009). Protocadherin-19 is essential for early steps in brain morphogenesis. Developmental Biology 334, 72–83.

Esumi S, Kakazu N, Taguchi Y, et al (2005) Monoallelic yet combinatorial expression of variable exons of the protocadherin-alpha gene cluster in single neurons. Nature genetics 37:171–176. doi: 10.1038/ng1500

Esumi S, Kaneko R, Kawamura Y, Yagi T (2006) Split single-cell RT-PCR analysis of Purkinje cells. Nature protocols 1:2143–2151. doi: 10.1038/nprot.2006.343

Etzrodt J, Krishna-K K, Redies C (2009) Expression of classic cadherins and δ-protocadherins in the developing ferret retina. BMC Neuroscience 10:153. doi: 10.1186/1471-2202-10-153

Geiss GK, Bumgarner RE, Birditt B, et al (2008) Direct multiplexed measurement of gene expression with color-coded probe pairs. Nature biotechnology 26:317–325. doi: 10.1038/nbt1385

Ghazanfar S, Bisogni AJ, Ormerod JT, et al (2016) Integrated single cell data analysis reveals cell specific networks and novel coactivation markers. BMC systems biology 10:127. doi: 10.1186/s12918-016-0370-4

Gierahn TM, Wadsworth MH, Hughes TK, et al (2017) Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. Nature methods 14:395–398. doi: 10.1038/nmeth.4179

Guo Y, Monahan K, Wu H, et al (2012) CTCF/cohesin-mediated DNA looping is required for protocadherin α promoter choice. Proceedings of the National Academy of Sciences of the United States of America 109:21081–21086. doi: 10.1073/pnas.1219280110

Hanchate NK, Kondoh K, Lu Z, et al (2015) Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. Science 350:1251–1255. doi: 10.1126/science.aad2456

Hertel N, Krishna-K, Nuernberger M, Redies C (2008) A cadherin-based code for the divisions of the mouse basal ganglia. The Journal of comparative neurology 508:511–528. doi: 10.1002/cne.21696

Hertel N, Redies C, Medina L (2012) Cadherin expression delineates the divisions of the postnatal and adult mouse amygdala. The Journal of comparative neurology 520:3982–4012. doi: 10.1002/cne.23140

Hirano K, Kaneko R, Izawa T, Kawaguchi M, Kitsukawa T, and Yagi T (2012). Single-neuron diversity generated by Protocadherin-β cluster in mouse central and peripheral nervous systems. Front Mol Neurosci 5, 90.

Horikoshi T, Sakakibara M (2000) Quantification of relative mRNA expression in the rat brain using simple RT-PCR and ethidium bromide staining. Journal of Neuroscience Methods 99:45–51.

Islam S, Kjällquist U, Moliner A, et al (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. Genome research 21:1160–1167. doi: 10.1101/gr.110882.110

Islam S, Zeisel A, Joost S, et al (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nature 11:163–166. doi: 10.1038/nmeth.2772

Kaneko R, Kato H, Kawamura Y, et al (2006) Allelic gene regulation of Pcdh-alpha and Pcdh-gamma clusters involving both monoallelic and biallelic expression in single Purkinje cells. Journal of Biological Chemistry 281:30551–30560. doi: 10.1074/jbc.M605677200

Kawasaki ES (2004) Microarrays and the gene expression profile of a single cell. Annals of the New York Academy of Sciences 1020:92–100. doi: 10.1196/annals.1310.010

Kehayova P, Monahan K, Chen W, Maniatis T (2011) Regulatory elements required for the activation and repression of the protocadherin-alpha gene cluster. Proceedings of the National Academy of Sciences of the United States of America 108:17195–17200. doi: 10.1073/pnas.1114357108

Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, and Taipale J (2012). Counting absolute numbers of molecules using unique molecular identifiers. Nat. Methods 9, 72–74.

Kolodziejczyk AA, Kim JK, Svensson V (2015) The Technology and Biology of Single-Cell RNA Sequencing. Molecular cell 58:610–620. doi: 10.1016/j.molcel.2015.04.005

Krishna-K K, Hertel N, and Redies C (2011). Cadherin expression in the somatosensory cortex: evidence for a combinatorial molecular code at the single-cell level. Neuroscience 175, 37–48.

Krishna-K K, Nuernberger M, Weth F, and Redies C (2009). Layer-specific expression of multiple cadherins in the developing visual cortex (V1) of the ferret. Cereb. Cortex 19, 388–401.

Kung JTY, Colognori D, Lee JT (2013) Long noncoding RNAs: past, present, and future. Genetics 193:651–669. doi: 10.1534/genetics.112.146704

La Manno G, Gyllborg D, Codeluppi S, et al (2016) Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. Cell 167:566–580. doi: 10.1016/j.cell.2016.09.027

Lin DM, Loveall B, Ewer J, et al (2007) Characterization of mRNA expression in single neurons. Methods in molecular biology 399:133–152.

Lin J, Wang C, Redies C (2012) Expression of δ-protocadherins in the spinal cord of the chicken embryo. The Journal of comparative neurology 520:1509–1531. doi: 10.1002/cne.22808

Luckner R, Obst-Pernberg K, Hirano S, et al (2001) Granule cell raphes in the developing mouse cerebellum. Cell and tissue research 303:159–172. doi: 10.1007/s004410000292

Luo R, Sanders SJ, Tian Y, et al (2012) Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders. American journal of human genetics 91:38–55. doi: 10.1016/j.ajhg.2012.05.011

Mackler SA, Eberwine JH (1993) Diversity of glutamate receptor subunit mRNA expression within live hippocampal CA1 neurons. Molecular pharmacology 44:308–315.

Macosko EZ, Basu A, Satija R, et al (2015) Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell 161:1202–1214. doi: 10.1016/j.cell.2015.05.002

Malkov VA, Serikawa KA, Balantac N, et al (2009) Multiplexed measurements of gene signatures in different analytes using the Nanostring nCounter™ Assay System. BMC Research Notes 2:80. doi: 10.1186/1756-0500-2-80

Marinov GK, Williams BA, McCue K, et al (2014) From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. Genome research 24:496–510. doi: 10.1101/gr.161034.113

Miyashiro K, Dichter M, Eberwine J (1994) On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. Proceedings of the National Academy of Sciences 91:10800–10804.

Monahan K, Rudnick ND, Kehayova PD, et al (2012) Role of CCCTC binding factor (CTCF) and cohesin in the generation of single-cell diversity of protocadherin-α gene expression. Proceedings of the National Academy of Sciences of the United States of America 109:9125–9130. doi: 10.1073/pnas.1205074109

Morrow EM, Yoo SY, Flavell SW, Kim TK, Lin Y, Hill RS, Mukaddes NM, Balkhy S, Gascon G, Hashmi A, et al. (2008). Identifying Autism Loci and Genes by Tracing Recent Shared Ancestry. Science 321, 218–223.

Mountoufaris G, Chen WV, Hirabayashi Y, et al (2017) Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. Science 356:411–414. doi: 10.1126/science.aai8801

Navin N, Kendall J, Troge J, et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472:90–94. doi: 10.1038/nature09807

Obst-Pernberg, K., and Redies, C. (1999). Cadherins and synaptic specificity. J. Neurosci. Res. 58, 130–138.

Patel AP, Tirosh I, Trombetta JJ, et al (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344:1396–1401. doi: 10.1126/science.1254257

Petropoulos S, Edsg rd D, Reinius B rn, et al (2016) Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. Cell 165:1012–1026. doi: 10.1016/j.cell.2016.03.023

Picelli S, Björklund ÅK, Faridani OR, et al (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nature methods 10:1096–1098. doi: 10.1038/nmeth.2639

Ramsköld D, Luo S, Wang Y-C, et al (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nature biotechnology 30:777–782. doi: 10.1038/nbt.2282

Redies C, Vanhalst K, and Roy FV (2005). delta-Protocadherins: unique structures and functions. Cell. Mol. Life Sci. 62, 2840–2852.

Rhee J, Buchan T, Zukerberg L, et al (2007) Cables links Robo-bound Abl kinase to N-cadherin-bound beta-catenin to mediate Slit-induced modulation of adhesion and transcription. Nature Cell Biology 9:883–892. doi: 10.1038/ncb1614

Ribich S, Tasic B, Maniatis T (2006) Identification of long-range regulatory elements in the protocadherin-alpha gene cluster. Proceedings of the National Academy of Sciences of the United States of America 103:19719–19724. doi: 10.1073/pnas.0609445104

Rodriguez S, Sickles HM, DeLeonardis C, et al (2008) Notch2 is required for maintaining sustentacular cell function in the adult mouse main olfactory epithelium. Developmental biology 314:40–58. doi: 10.1016/j.ydbio.2007.10.056

Roozemond RC (1976) Ultramicrochemical determination of nucleic acids in individual cells using the Zeiss UMSP-I microspectrophotometer. Application to isolated rat hepatocytes of different ploidy classes. The Histochemical Journal 8:625–638. doi: 10.1007/BF01003963

Ross EM, Markowetz F (2016) OncoNEM: inferring tumor evolution from single-cell sequencing data. Genome Biology 17:69. doi: 10.1186/s13059-016-0929-9

Saraiva LR, Ibarra-Soria X, Khan M, et al (2015) Hierarchical deconstruction of mouse olfactory sensory neurons: from whole mucosa to single-cell RNA-seq. Nature Publishing Group 5:18178. doi: 10.1038/srep18178

Schneider M, Huang C, Becker SFS, et al (2014) Protocadherin PAPC is expressed in the CNC and can compensate for the loss of PCNS. Genesis (New York, NY : 2000) 52:120–126. doi: 10.1002/dvg.22736

Shapiro L, and Colman DR (1999). The diversity of cadherins and implications for a synaptic adhesive code in the CNS. Neuron *23*, 427–430.

Shekhar K, Lapan SW, Whitney IE, et al (2016) Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. Cell 166:1308–1323. doi: 10.1016/j.cell.2016.07.054

Stevens A, Jacobs JR (2002) Integrins regulate responsiveness to slit repellent signals. The Journal of neuroscience : the official journal of the Society for Neuroscience 22:4448–4455.

Sucher NJ, Deitcher DL, Baro DJ, et al (2000) Genes and channels: patch/voltage-clamp analysis and single-cell RT-PCR. Cell and tissue research 302:295–307. doi: 10.1007/s004410000289

Svensson V, Natarajan KN, Ly L-H, et al (2017) Power analysis of single-cell RNA-sequencing experiments. Nature methods. doi: 10.1038/nmeth.4220

Tang F, Barbacioru C, Wang Y, et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nature methods 6:377–382. doi: 10.1038/nmeth.1315

Tan L, Li Q, Xie XS (2015) Olfactory sensory neurons transiently express multiple olfactory receptors during development. Molecular systems biology 11:844. doi: 10.15252/msb.20156639

Tasic B, Menon V, Nguyen TN, et al (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nature neuroscience 19:335–346. doi: 10.1038/nn.4216

Tasic B, Nabholz CE, Baldwin KK, et al (2002) Promoter choice determines splice site selection in protocadherin alpha and gamma pre-mRNA splicing. Molecular Cell 10:21–33.

Toledo-Rodriguez M, Markram H (2007) Single-cell RT-PCR, a technique to decipher the electrical, anatomical, and genetic determinants of neuronal diversity. Methods in molecular biology (Clifton, NJ) 403:123–139. doi: 10.1007/978-1-59745-529-9_8

Tsai N-P, Wilkerson JR, Guo W, et al (2012) Multiple autism-linked genes mediate synapse elimination via proteasomal degradation of a synaptic scaffold PSD-95. Cell 151:1581–1594. doi: 10.1016/j.cell.2012.11.040

Uemura E (1980) Age-related changes in neuronal RNA content in rhesus monkeys (Macaca mulatta). Brain Research Bulletin 5:117–119. doi: 10.1016/0361-9230(80)90182-3

Uemura M, Nakao S, Suzuki ST, Takeichi M, and Hirano S (2007). OL-Protocadherin is essential for growth of striatal axons and thalamocortical projections. Nat. Neurosci. 10, 1151–1159.

Vanhalst K, Kools P, Staes K, van Roy F, and Redies C (2005). delta-Protocadherins: a gene family expressed differentially in the mouse brain. Cell. Mol. Life Sci. 62, 1247–1259.

Wang X, Su H, Bradley A (2002) Molecular mechanisms governing Pcdh-gamma gene expression: evidence for a multiple promoter and cis-alternative splicing model. Genes & development 16:1890–1905. doi: 10.1101/gad.1004802

Wang Y, Waters J, Leung ML, et al (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512:155–160. doi: 10.1038/nature13600

Wu C, Niu L, Yan Z, et al (2015) Pcdh11x Negatively Regulates Dendritic Branching. Journal of Molecular Neuroscience 56:822–828. doi: 10.1007/s12031-015-0515-8

Yagi T (2012) Molecular codes for neuronal individuality and cell assembly in the brain. Frontiers in molecular neuroscience 5:45. doi: 10.3389/fnmol.2012.00045

Yamagata K, Andreasson KI, Sugiura H, et al (1999) Arcadlin is a neural activity-regulated cadherin involved in long term potentiation. Journal of Biological Chemistry 274:19473.

Yasuda S, Tanaka H, Sugiura H, et al (2007) Activity-Induced Protocadherin Arcadlin Regulates Dendritic Spine Number by Triggering N-Cadherin Endocytosis via TAO2β and p38 MAP Kinases. Neuron 56:456–471. doi: 10.1016/j.neuron.2007.08.020

Chapter 3

The Adhesive Functions of δ-Protocadherins in Combination

1. Abstract

Cell adhesion molecules are critical for proper neural circuit formation during development. In order to meet the demands of the nervous system, cell adhesion molecules such as the δ-protocadherins have been proposed to generated combinatorial adhesive codes to specify neural wiring choices. In Chapter 2, we demonstrated that δ-protocadherins are expressed in combinations within single neurons, confirming that their expression patterns can generate combinatorial codes. However, the deltas have been poorly studied individually and never studied as a family. Here, we studied the adhesive functions of deltas individually, and in simple combination. We show that the deltas display a range of apparent affinities, allowing some to bind stronger and overcome mismatched deltas when coexpressed. We propose that combinatorial function of deltas ultimately depends on which deltas are expressed, how many are expressed, and their relative surface expression levels.

2. Introduction

The formation of the nervous system has long been theorized to require combinations of guidance cues to work together to enable billions of neurons to reach and synapse with their appropriate partner (Sperry 1963; Shapiro and Colman 1999). But while genetic, *in vitro*, and genomic approaches have revealed how any one cue can influence this process, how these many members work together to mediate guidance and target recognition is still very poorly understood. One major class of guidance cues thought to work in combination are cell adhesion molecules (CAMs). Although CAMs themselves are highly diverse, expression studies have shown that individual neurons are likely to express different combinations of CAMs. For example, the cadherins are the largest category of CAMs (Hulpiau et al. 2016)**,** and *in situ* hybridization studies have suggested that individual neurons clearly express different combinations of CAMs (Wöhrn et al. 1999; Redies 2003; Krishna-K et al. 2011). But what the full extent of these combinations are, and how they work together to exert an effect on neuronal behavior is still unknown.

Several approaches have been used to understand how combinations of cues mediate guidance and synapse formation. Genetic tools have been used to affect the expression of two, and sometimes three, different cues at one time. However, this approach is time consuming and relatively difficult in vertebrates. *In vitro* approaches have utilized primary neuronal culture to study the impact of CAMs and other cues on behavior. However, such studies by their nature utilize a heterogeneous population of neurons,

and require a population-based approach to understand how cues may function together. Structural analyses can begin to show how cues may physically interact. This in turn can generate hypotheses for further study, but do not in and of themselves demonstrate function. Finally, the advent of single-cell genomic tools can, for the first time, reveal the particular combination of cues expressed by individual neurons. But such approaches still do not provide a functional readout to determine how CAMs and other cues affect neuronal function.

To begin to address the complexity associated with understanding combinatorial CAM function, we have chosen to study the δ-protocadherins, a nine member subfamily of the cadherin superfamily. δ-Protocadherins are known cell adhesion molecules (Sano et al. 1993; Hirano et al. 1999; Tai et al. 2010; Izuta et al. 2015), and genetic deletion and misexpression studies have shown a role for individual deltas in axon guidance and synapse formation (Uemura et al. 2007; Williams et al. 2011; Emond et al. 2011; Hayashi et al. 2014). In Chapter 2, we showed that randomly selected olfactory sensory neurons can express anywhere from zero to seven different deltas, consistent with *in situ* patterns throughout the nervous system (Vanhalst et al. 2005; Krishna-K et al. 2011; Lin et al. 2012a; Lin et al. 2012b). To determine the functional effect of these combinations on cellular behavior, we utilized the K562 adhesion assay to study how different combinations affected cell adhesion. Finally, we propose a working model to explain how combinatorial delta expression regulates adhesive cellular behaviors.

3. Methods

3.1. DNA Plasmids

All constructs were cloned into modified pEGFP-N1 vectors (p2a and direct fusion) with protocadherin fragments generated by PCR from the ATG to the end of the transmembrane domain for each delta with either eGFP, TagRFP, or FLAG tags at the C-terminus.

3.2. K562 Aggregation Assay

K562 myelogenous leukemia cells were obtained from ATCC (ATCC CCL-243) and maintained in RPMI + L-glutamine with 10% calf bovine serum. To transfect δ-protocadherins into K562 cells, electroporations were performed using an Amaxa Nucleofector II (Lonza) when cells reached a density between 250-500,000 cells/mL. For each reaction, 1 x 10^6 cells were gently pelleted and then resuspended in 100 µL of Ingenio Electroporation Solution (Mirus Bio), and electroporated using program T-016 with 5-8ug of DNA, and immediately recovered in 2 mL of incubator equilibrated media. After 1 hour, 4mM valproic acid was added to the cultures to promote maintained expression, and in experiments where two populations were mixed, equal volumes from each population were then combined. Both 6 well and 24 well plates were used, with total volumes of 2mL and 500 respectively. Cells were gently agitated continuously at 15 RPM overnight in a tissue culture incubator, and images were taken between 24-26 hours post-electroporation. Most images were captured using confocal microscopy at 1024 x 1024 using a 10x objective (Zeiss LSM 510), although some images were

captured using a camera attached to an inverted fluorescent scope with a 10x or 20x objective.

3.3. Coaggregation Index Analysis

To generate the Coaggregation Index, confocal images were analyzed using custom code written in Mathematica (Wolfram Research). Briefly, each confocal image is divided into squares with areas slightly larger than the area of one cell (15 x 15 pixels) to allow for sampling of neighboring cells. Squares with no fluorescent signal (all black) are discarded, and the remaining squares are analyzed for how many colors are detected in each square. The final output is the percent of squares with colors that have more than one color, which we termed the CoAggregation Index.

3.4. Aggregate Size Analysis

For the initial assessment of δ-protocadherin mediated aggregation size, two independent electroporations were performed. For each replicate, 15-20 fields of view were imaged using an inverted fluorescent microscope with a 10x objective. These images were then analyzed in ImageJ to measure aggregate size. Pixel size of aggregates were compared to the pixel area of one cell to approximate the number of cells per aggregate. Aggregates smaller than three cells were removed from the analysis to prevent dividing cells and single cells not participating in aggregation from skewing the results.

3.5. Aggregate Size with Increased Rotational Speeds

K562 cells were electroporated and allowed to form aggregates at 15 RPM overnight. At 24 hours, six fields of view were captured with confocal microscopy with a 5x objective, and then returned to the incubator. The speed was then increased to 120 RPM for 1 hour, and images were captured again. This process was repeated at 160, 200 and 220 RPM for a total of five time points. Three independent electroporations were performed for each sample. Each image was then analyzed using a custom written code in Mathematica (Wolfram Research) to measure the pixel size of each aggregate, and aggregate pixel size was converted to microns.

3.6. Surface Biotinylation

Surface biotinylation of live K562 cells was performed using the Pierce Cell Surface Isolation Kit (Thermo Scientific), and the protocol was modified for use with smaller volumes and smaller columns (Pierce Micro Spin Columns, Thermo Scientific). Briefly, at 24 hours post-electroporation, cells were gently pelleted at ~ 1000 x g, and then washed once with ice cold PBS. The labeling solution was prepared by suspending 12 mg of EZ-Link Sulfo-NHS-SS-Biotin in 48 mL of ice cold PBS. Cell pellets were then resuspended into single cell suspensions with 1 mL of the labeling solution, and moderately agitated at 4C on a rotary shaker such that the solution was constantly mixing for 30 minutes. Next, 50 µL of quenching solution was added to each sample, and the cells were again pelleted at 1000 x g for 3 minutes, and washed once with 1 mL ice cold TBS. The pellets were then lysed on ice for 30 minutes using 50 µL of the

supplied lysis buffer plus added protease inhibitors. The lysate was then spun at 10,000 x g for 10 minutes at 4C, and the supernatant was transferred to micro columns containing 100 µL of Neutravidin agarose slurry that had been washed four times with 100 µL of wash buffer by spinning at 1000 x g for 1 minute at room temperature. An additional 150 µL of lysis buffer was added to ensure complete mixing during incubation. Samples were then capped and incubated for 1 hour at room temperature with end-over-end rotary mixing. The columns were then spun at 1000 x g for 1 minute, and washed with 4 x 100 µL wash buffer with protease inhibitors. Finally, 50 µL of SDS-Page sample buffer with 50 mM DTT was added to each column, heated for 5 minutes at 95C, and then spun for 3 minutes at 20,000 x g to elute the purified sample, and 1 µL of bromophenol blue was added to each sample.

3.7. Western Blot Analysis

Western blots were performed by loading 8 µL (roughly 15% of total elution from each biotinylation experiment) onto 10% SDS polyacrylamide gels. The separated proteins were then transferred to polyvinylidene difluoride (PVDF) membranes for 2.5 hours at 380 mAmps at 4C. The membranes were then blocked in 5% non-fat dairy milk in TBST (Tris-buffered saline with 0.1% Tween 20) for 1 hour at room temperature, and then incubated in primary antibodies overnight for 16 to 18 hours. All primary antibodies used were monoclonal, and carefully titrated to establish working dilutions of equivalent detection so that samples across antibodies could be compared. To achieve this, whole cell lysates of K562 cells electroporated with equal amounts of DNA for Pcdh7 fused with GFP, RFP and FLAG were used. Equal amounts of lysate were probed at various

concentrations for each antibody, and over several rounds of optimization, dilutions for each antibody that produced roughly equivalent detection were obtained. The antibodies used were mouse anti-GFP (1:4,000, Thermo Scientific MA5-15256), mouse anti-RFP (1:2,000, Thermo Scientific MA5-15257) and mouse anti-FLAG (1:6,000, Thermo Scientific MA1-91878). The best loading control for surface expression was determined to be transferrin receptor (TfR) and was used at 1:1,000 (Thermo Scientific 13-6800). After primary incubation, blots were washed 3 x TBST for 5 minutes, then incubated with secondary goat anti-mouse HRP antibody (1:10,000, Thermo Scientific) for 1.5 hours, wash 3 x TBST for 5 minutes again, and then developed with ECL. Estimation of band intensity was carried out using ImageJ.

4. Results

4.1. δ-Protocadherins are Calcium Dependent Homophilic Cell Adhesion Molecules

A subset of δ-protocadherins have been tested for their adhesive abilities using a variety of different approaches across different studies (Sano et al. 1993; Hirano et al. 1999; Yoshida 2003; Tai et al. 2010; Emond et al. 2011; Izuta et al. 2015), but have never been systematically characterized as a family. We assessed the cell adhesion properties of the deltas using the K562 cell adhesion assay. K562 cells are excellent cells to test for cell-adhesion because they do not endogenously express any cadherins or exhibit any cell-cell adhesion (Schreiner and Weiner 2010). Thus, any cell aggregation behaviors can be attributed to what they were manipulated to express.

Briefly, K562 suspension cells are electroporated with DNA constructs encoding δ-protocadherins tagged with GFP or RFP. The cells are allowed to recover for 30-60 minutes, and are then gently agitated at 15 RPM and observed at 24 hours post electroporation. Expression of full-length constructs was more difficult than only expressing the extracellular and transmembrane (ECTM) domains. However, preliminary tests comparing full-length Pcdh1 against just the ECTM of Pcdh1 produced equivalent aggregates, and the ECTM was able to bind homophilically (binding with like protocadherin on opposing cell) to the full-length construct suggesting that the lack of the intracellular domain (ICD) did not affect adhesive recognition outside the cell (Figure 3.1). This suggests that the intracellular domain is not required to mediate adhesive events. This is consistent with a few other reports on protocadherins that also removed the intracellular domain without disrupting adhesion (Chen et al. 2007; Tai et al. 2010; Emond et al. 2011). Therefore, ECTM constructs were used in all subsequent experiments.

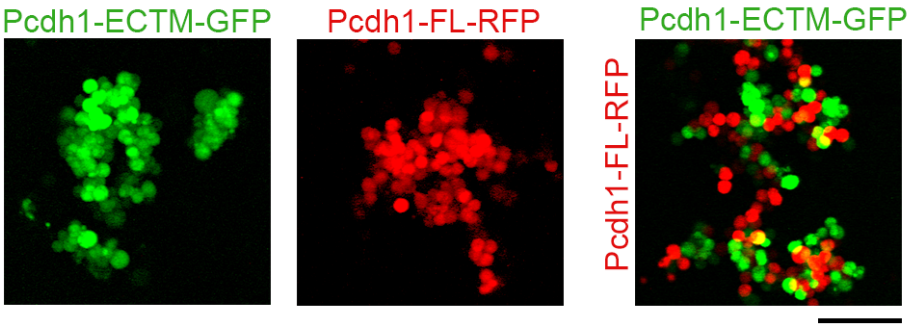

Figure 3.1 The δ-protocadherins still exhibit homophilic recognition, even when one population lacks the intracellular domain. (**Left**) A typical Pcdh1-ECTM-GFP aggregate. (**Middle**) A typical Pcdh1-FL-RFP aggregate. (**Right**) Both populations engage in highly mixed homophilic aggregates when allowed to co-mingle. Scale bar =100 µm.

We then tested eight of the nine deltas in this format (Pcdh11x did not express in K562 cells) for their ability to induce aggregation. GFP alone does not mediate any aggregation, and cells remain in a single cell suspension. Interestingly, the eight deltas all mediated adhesion, but to varying degrees (Figure 3.2A). For example, Pcdh10 formed very small aggregates relative to the other deltas. Varying the amount of Pcdh10 transfected into cells had a limited impact on the size of aggregates formed. Since it induced significantly smaller aggregates than that seen with the other deltas, Pcdh10 was excluded from subsequent experiments. In contrast, K562 cells transfected with Pcdh7 or Pcdh17 formed larger aggregates than the other deltas. In order to normalize the size of aggregates, the amount of DNA electroporated for each delta was titrated to generate an approximately equal distribution of aggregate sizes (Figure 3.2B). The importance of modulating the level of expression of a given delta in relation to adhesive function was further examined in later studies.

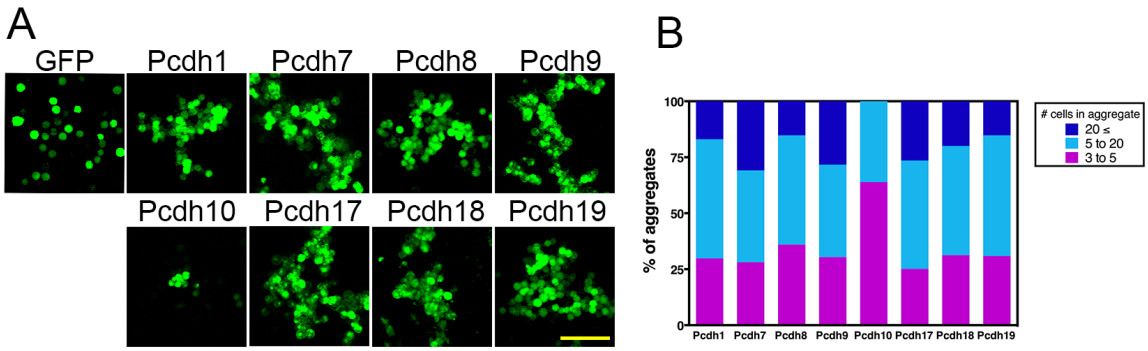

Figure 3.2 (**A**) Electroporation of δ-protocadherins into K562 cells induces aggregation formation, whereas GFP alone does not. (**B**) Distribution of aggregate sizes for each delta from two independent electroporations**.** Scale Bar = 100 µm.

.

A defining feature of all cadherin superfamily members is that their adhesive functions are calcium dependent. Presumably, this is one way in which their function may be

regulated, although how extracellular calcium dynamics influence cadherin function is not well understood. The δ-protocadherins have never been systematically characterized for this defining cadherin characteristic. We therefore next asked if δ-protocadherin cell adhesion and aggregate formation was calcium-dependent by comparing aggregate formation of each delta in normal media (RPMI 1640, containing 42 μM $CaCl_2$), and media with 20 μM EDTA, which should be sufficient to chelate all calcium ions present in RPMI1640. Control cells electroporated with only GFP remained as single cell suspensions with both treatments (Figure 3.3). Consistent with being members of the cadherin superfamily, all deltas had significantly reduced levels of aggregation in the presence of EDTA. However, some deltas may be less sensitive to calcium than others, as some (e.g. Pcdh7, Pcdh8, and Pcd17) still maintained aggregates, albeit small ones. The addition of higher levels of $CaCl_2$ in the media did not have a noticeable effect on cell aggregation and aggregate size (data not shown), suggesting that they are already saturated with calcium at 42 μM. This suggests that the deltas are largely calcium dependent, but some calcium independent adhesion may still occur for at least some members. The significance of these slight differences will need to be addressed in future studies.

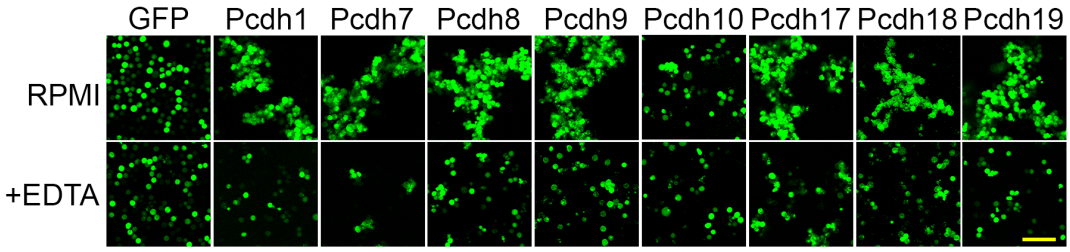

Figure 3.3 δ-protocadherin aggregation is primarily calcium dependent, as aggregation is severely disrupted by the presence of 20 μM EDTA. However, some deltas still displayed small aggregates in the presence of EDTA (e.g. Pcdh8 and Pcdh17). Scale bar =100 μm.

4.2. δ-Protocadherin Binding in *trans* is Strictly Homophilic

To determine if deltas mediate only homophilic aggregation, we next mixed cells expressing a given delta (fused to p2a-GFP) with those expressing another delta (fused to p2a-RFP) and tested all 49 possible pairs. As expected, RFP and GFP cells expressing the same delta mixed homophilically (Figure 3.4, squares along center diagonal), while those expressing different deltas segregated from one another and formed distinct and completely separate aggregates (Figure 3.4, all squares off center diagonal). These results suggest that in *trans,* δ-protocadherins will preferentially bind with the same δ-protocadherin. Therefore, we conclude that δ-protocadherins are strictly homophilic when binding in *trans.*

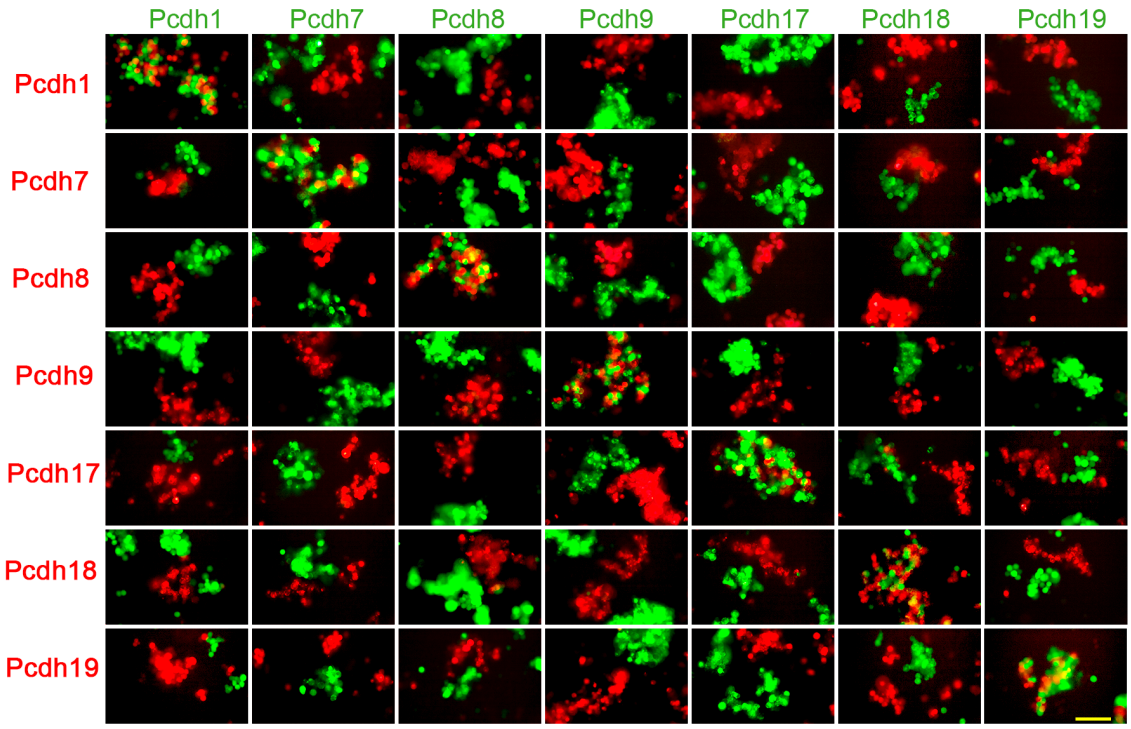

Figure 3.4 δ-Protocadherins exhibit strict homophilic adhesion in *trans*. Cells expressing deltas with a p2a-GFP tag (columns) were mixed with cells expressing deltas with p2a-RFP tags. All possible pair-wise combinations were tested. Each pair was tested in two independent electroporations. Scale bar =100 µm.

4.3. Mismatched Delta Populations Generate Various Coaggregation Behaviors

The results shown in Figure 4 are analogous to those previously found using clustered protocadherins (Schreiner and Weiner 2010; Thu et al. 2014). In those studies, further experiments were also performed whereby the protocadherins expressed between the two populations were “mismatched”. In the simplest example, cells expressing one clustered protocadherin were allowed to mix with cells expressing the same clustered

protocadherin plus an additional, different clustered protocadherin. In all variations tested in the K562 assay, the two populations always segregated from one another, suggesting that a single mismatch was sufficient to prevent any mixing. To determine if the δ-protocadherins followed the same general principle, we performed similar experiments.

In these experiments, we systematically "challenged" cells expressing just one δ-protocadherin to coaggregate with cells expressing the same delta plus an additional "mismatched" delta. We therefore refer to these experiments as "mismatch coaggregation assays" (Figure 3.5A). We tested all 42 possible mismatch pairs and critically, we discovered that the uniform principle defined for the clustered protocadherins does not appear to apply to the deltas (Figure 3.5B, C). Instead, we observed a range of outcomes which can be described as three general categories of aggregation behavior. In the first, we found the populations completely segregated from one another, as seen with the clustered protocadherins (Figure 3.5B, Pcdh7+Pcd19 vs Pcdh19). In the second, the populations appeared to "interface," as red aggregates directly abutted green aggregates (Figure 3.5B, Pcdh1+Pcdh7 vs Pcdh1). In the last category, green and red cells appeared to intermix freely with one another (Figure 3.5B, Pcdh7+Pcdh19 vs Pcdh7). Importantly, however, the extent of interfacing or intermixing appeared to vary depending on the particular deltas used in the experiment.

In previous coaggregation experiments performed by others, outcomes were primarily considered either binary (mixing or no mixing) or placed into categories (e.g. type I, II, III, etc.) by eye to interpret different outcomes. We choose to quantify our observations

by developing a novel analysis to produce a metric (CoAggregation Index) of how well two different populations coaggregate. This allowed us to detect and quantify subtle differences without bias. Briefly, this metric (see methods) parses an image into a grid where the area of each square is just slightly larger than the size of one K562 cell. Each square of the grid is then analyzed for the presence of red, green, or yellow signal. To determine the extent of mixing, the number of squares containing more than one color are counted and reported as a percent (corrected for squares without any signal). Cells with strong intermixing will have a higher proportion of squares containing more than one color, generating a higher CoAgg Index value.

Analyzing these initial results with this metric allowed us to identify more nuanced differences among the three broad phenotypic categories. Indeed, subtle differences not visible by eye could be easily distinguished as they produced different CoAgg Index values (Figure 3.5B). Applying our quantification method to all 42 mismatch coaggregation assays, we observed a broad spectrum of CoAgg Index values (Figure 3.5D). On the high end of the CoAgg Index, some mismatched pairs (e.g. Pcdh7 vs Pcdh7+Pcdh19, Figure 3.5B, first image) intermixed just as well as two populations expressing the same protocadherin without a mismatch (e.g. Pcdh7 vs Pcdh7, purple bar). On the opposite end, some did not mix at all (e.g. Pcdh19 vs Pcdh7+Pcdh19, Figure 3.5B, far right image), similar to control experiments with two different protocadherins (e.g. Pcdh1 vs Pcdh7, green bar). In the middle we observed a continuum of intermixing and interfacing. As a general rule, CoAgg values below 0.1 indicated segregation. From 0.1 to 0.2, clear interfacing is observed, and as 0.2 is

surpassed the behavior begins to transition from interfacing towards intermixing. Generally, CoAgg values above 0.5 were rarely observed. Thus, the use of the CoAgg Index revealed that interfacing and intermixing behaviors lie on a continuum. The use of our metric allows us to distinguish visually similar outcomes that would not be detectable by eye.

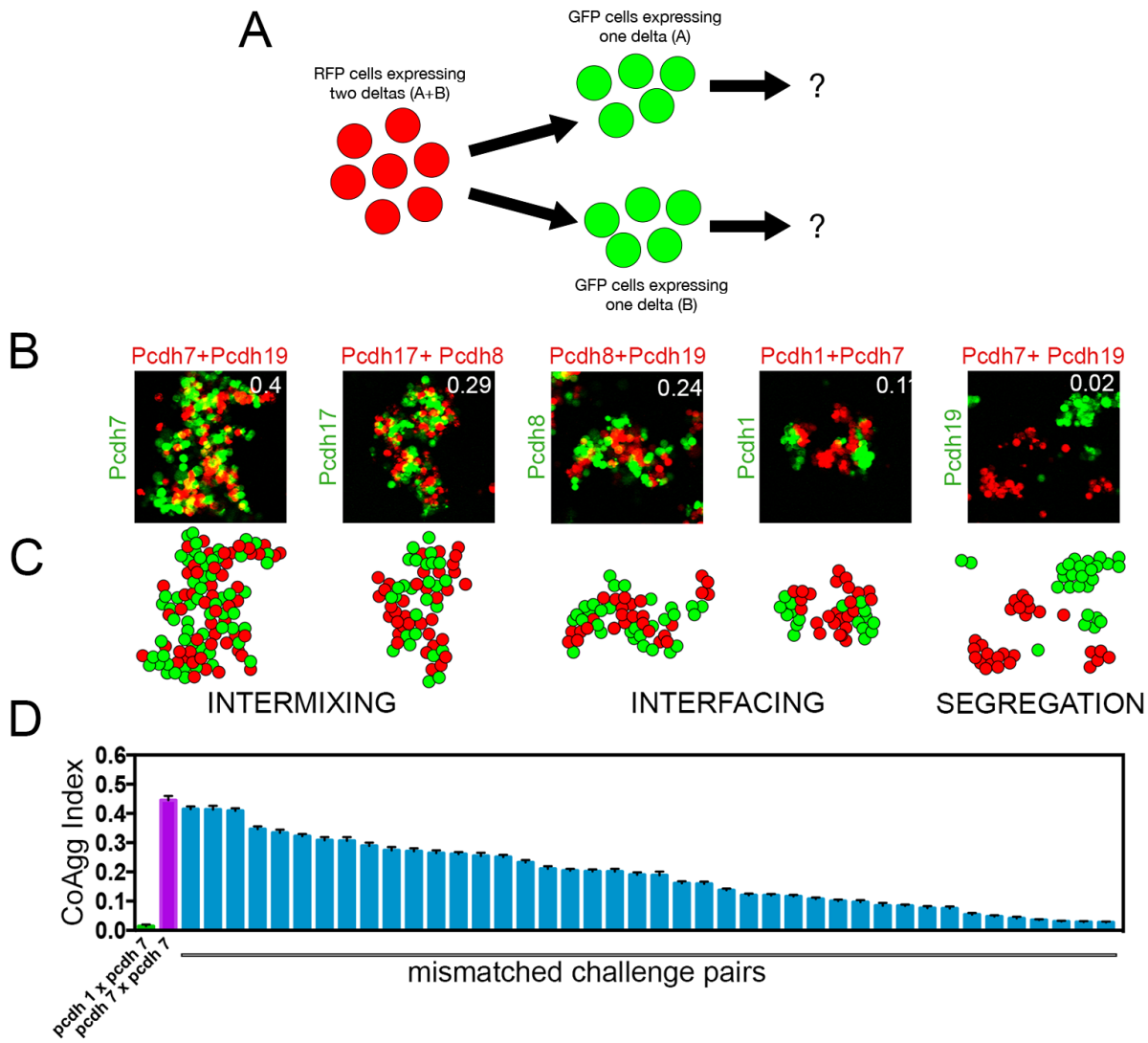

Figure 3.5 Spectrum of coaggregation behaviors observed in mismatch coaggregation assay screen. (**A**) Diagram of mismatch coaggregation assay, where the RFP population expresses two deltas, and each GFP population expresses only one of the two deltas. (**B**) Select pairs from screen displaying different possible outcomes along the spectrum. Number in upper right corner is average CoAgg Index value for that pair. (**C**) Graphical representations of aggregate formation in A. (**D**) Spectrum of CoAgg indexes measured in mismatch coaggregation assay, ordered from highest to lowest. Note the comparisons to the single pair-wise tests on the far left. Error bars indicate standard error. Each pair in the screen was performed over two independent electroporations. Scale bar = 100 µm.

.

A critical principle that emerged from these experiments was that the CoAgg Index depended on the identity of the mismatched protocadherin (Figure 3.6A). For example, for Pcdh1 (Figure 3.6A, row 1), the index was high when mixed with cells expressing Pcdh1+Pcdh9 (dark blue, index = 0.30), indicating strong intermixing. In contrast, the index was low when these cells were mixed with cells expressing Pcdh1+Pcdh7 (medium blue, index = 0.11), consistent with an interfacing behavior. Comparing the behavior of different deltas in these assays (columns) revealed different behaviors for each protocadherin. For example, for column 1, when Pcdh1+Pcdh7 cells are mixed with Pcdh7, the index value is 0.33 (dark blue, intermixing), whereas Pcdh1+Pcdh9 cells mixed with Pcdh9 cells gave an index of 0.07 (light blue, segregation).

In addition to these differences in any one row or column, we noted that differences existed for complementary pairs (Figure 3.6A, 6B). Notably, a heat map of these data displays great asymmetry across the diagonal. Thus, mixing Pcdh1+Pcdh7 cells with Pcdh1 cells (index of 0.11) resulted in interfacing, whereas mixing Pcdh1+7 cells with Pcdh7 cells resulted in strong intermixing, producing an index of 0.33. Although all deltas are adhesive (Figure 3.2), we conclude that their ability to interact with cells expressing distinct combinations of deltas is strongly dependent upon which deltas are expressed within a given combination.

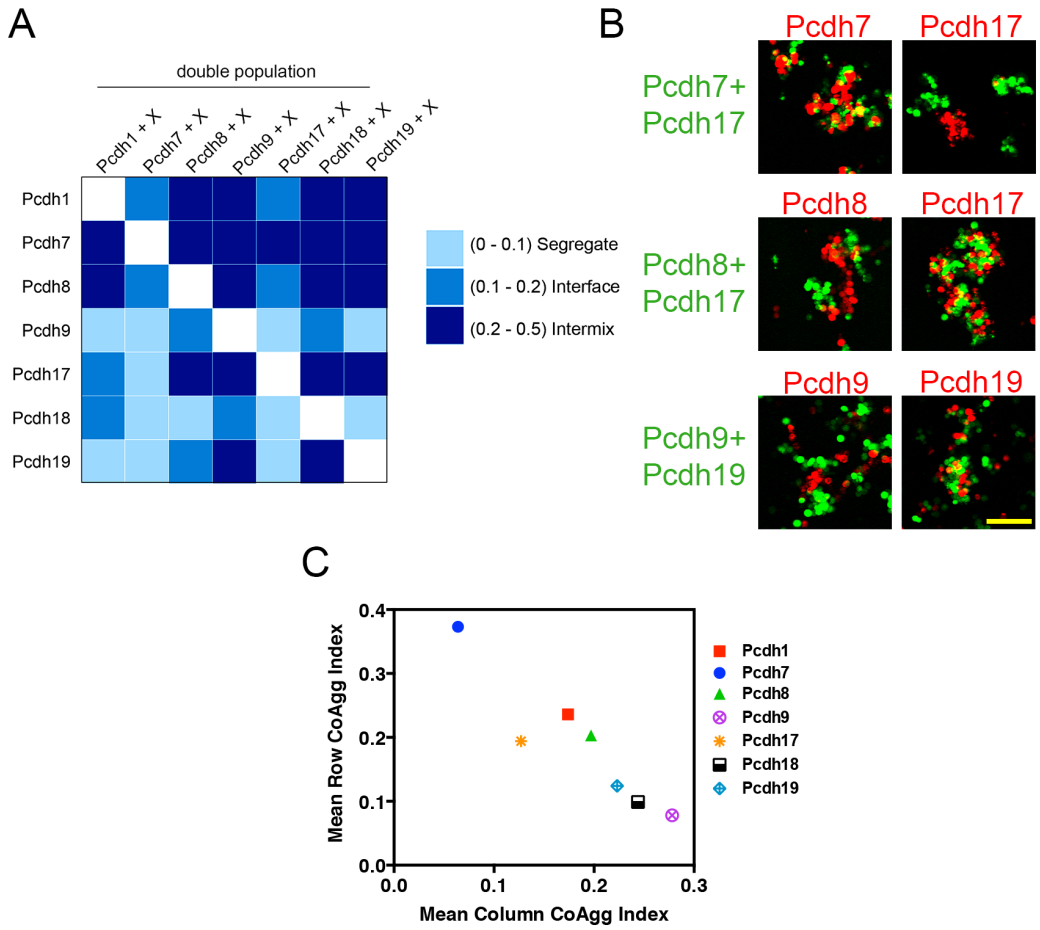

Figure 3.6 Differential behaviors in mismatch coaggregation assay. (**A**) Heat map of CoAgg Index values reveals lack of symmetry across pairs. (**B**) Examples of non-complementary results for mismatched pairs. (**C**) Estimated rankings of intermixing behaviors based on mismatch coaggregation assay screen mean row and column scores. Scale bar = 100 μm

Since there is a clear range in behavior among the deltas, we sought to estimate the overall ranking for how well each delta was able to 'overcome' the presence of a mismatch and still intermix. Using the CoAgg index values from our screen of 42 mismatch coaggregation assays, we estimated a hierarchy or ranking of each delta. To do this, we first calculated the mean CoAgg Index across the rows for each delta. This indicates how well each protocadherin alone can mix with cells that also express the

same protocadherin plus an additional, mismatched δ-protocadherin. In this case, higher index scores indicate a stronger ability to intermix with a population that expresses a mismatched delta. Next, the column means for each delta were calculated. In contrast to row means, lower column index values indicate stronger ability to prevent mixing when acting as the mismatched gene. Plotting each protocadherin with the row means on the Y-axis and the column means on the X-axis creates a linear trend and ranking of the deltas, suggesting these two parameters are capturing the same information. These results suggest that the deltas display a wide range in their abilities to overcome mismatched pairs, and some are better than others at overcoming mismatched pairs to still bind with their matched delta.

We hypothesized three potential explanations for the behavior of different deltas in our mismatched coaggregation assays. First, differences in the surface expression of each delta expressed within the various cell populations could influence their adhesive interactions. Thus, the expression of two deltas within a given cell may result in uneven (or non-existent) surface expression of one relative to the other, biasing the adhesive interactions. Second, different deltas may possess different apparent adhesive affinities. Stronger affinities would produce more stable binding events, and increase the probability of intermixing despite mismatched proteins being present. And finally, some combination of the two may occur. In this case, the ultimate adhesive behavior of a cell expressing two deltas would be determined by which deltas are expressed, the apparent affinities of each delta, and their relative expression levels.

4.4. Differential Coaggregation Outcomes are Not Explained by Differences in Surface Expression

To address the first possibility, we repeated select mismatch coaggregation experiments and controlled for surface protein expression of each delta. To track proteins at the surface, we generated ECTM constructs of each delta fused with FLAG, GFP, and RFP. Through several rounds of optimization, we carefully titrated the amount of each construct transfected into cells to minimize differences in surface protein expression, as confirmed using surface protein biotinylation and western blots. As a control, we showed that the biotinylation of the electroporated cells specifically detected surface protein expression and not cytosolic proteins (Figure 3.7)

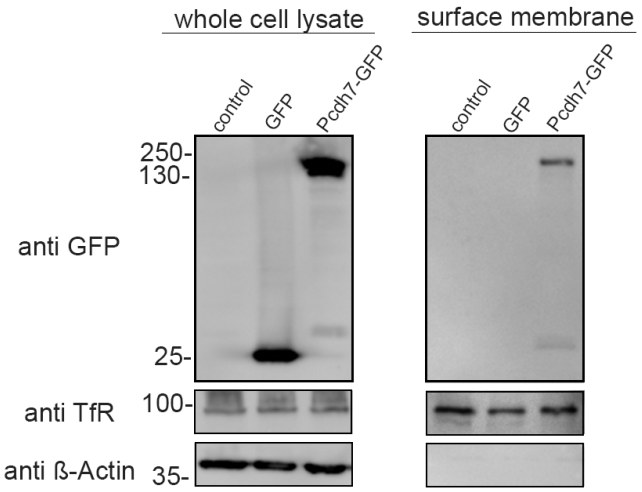

Figure 3.7 Surface biotinylation of K562 cell surface is specific to the plasma membrane and does not detect cytoplasmic proteins. Non-electroporated control cells, and cells electroplated with cytosolic GFP and Pcdh7 fused with GFP were used. (**Left**) Whole cell lysate shows no GFP signal in control cells, and strong GFP and Pcdh7-GFP cells. All samples had strong cytosolic signal for β-actin signal and membrane bound Transferrin receptor (TfR). (**Right**) In the biotinylated samples, no cytosolic β-actin was detected in any samples, and TfR was enriched. GFP antibodies only detected Pcdh7-GFP samples, suggesting only proteins on the cell surface were isolated.

For the first test, we chose Pcdh1 and Pcdh7, as these had different CoAgg Index values when challenged to coaggregate with cells expressing Pcdh1+Pcdh7 (0.11 and 0.33, respectively). After optimizing surface expression conditions such that each delta was present at the surface at roughly the same level (Figure 3.8A), we repeated the mismatched coaggregation assay, and imaged aggregate formation over several time points (Figure 3.8B,C). Consistent with our initial screen, and having controlled for differences in expression level, we still found that Pcdh7 cells consistently produced a higher index value when combined with cells expressing Pcdh1+Pcdh7. In contrast, a mismatch coaggregation assay where Pcdh1 cells were combined with cells expressing Pcdh1+Pcdh7 consistently resulted in lower values and only interfaced. We repeated this same experiment using Pcdh1+Pcdh17 to determine if this result was generalizable (Figure 3.8D). Since the Pcdh1+Pcdh7 time course indicated that 26 hours was sufficient to detect mature aggregates, we only captured the Pcdh1+Pcdh17 set at this time point (Figure 3.8E,F). As with Pcdh1+Pcdh7, Pcdh1+Pcdh17 also demonstrated differences in coaggregation behavior depending upon which delta was mismatched. While Pcdh17 expressing cells could strongly intermix with Pcdh1+Pcdh17 cells (index = 0.44), cells expressing Pcdh1 segregated (index = 0.04). We note, however, that this is a slight change from our initial screen, which found that cells expressing Pcdh1 and Pcdh17 alone both interfaced with cells co-expressing Pcdh1+Pcdh17. This suggests that some pairs in the screen might have been influenced by unequal surface expression. Interestingly, we also note that Pcdh17 is the only gene that is slightly off the linear trend in our hierarchical ranking of deltas from the mismatch coaggregation

assay screen (Figure 3.6C), suggesting that this may be the cause. However, the results when the levels are balanced are still consistent with different behaviors of mixing occur depending on which deltas are being tested. Therefore, it is unlikely that our first interpretation involving differential surface expression is correct. Instead, we interpret these results to indicate that different deltas possess different apparent adhesive affinities.

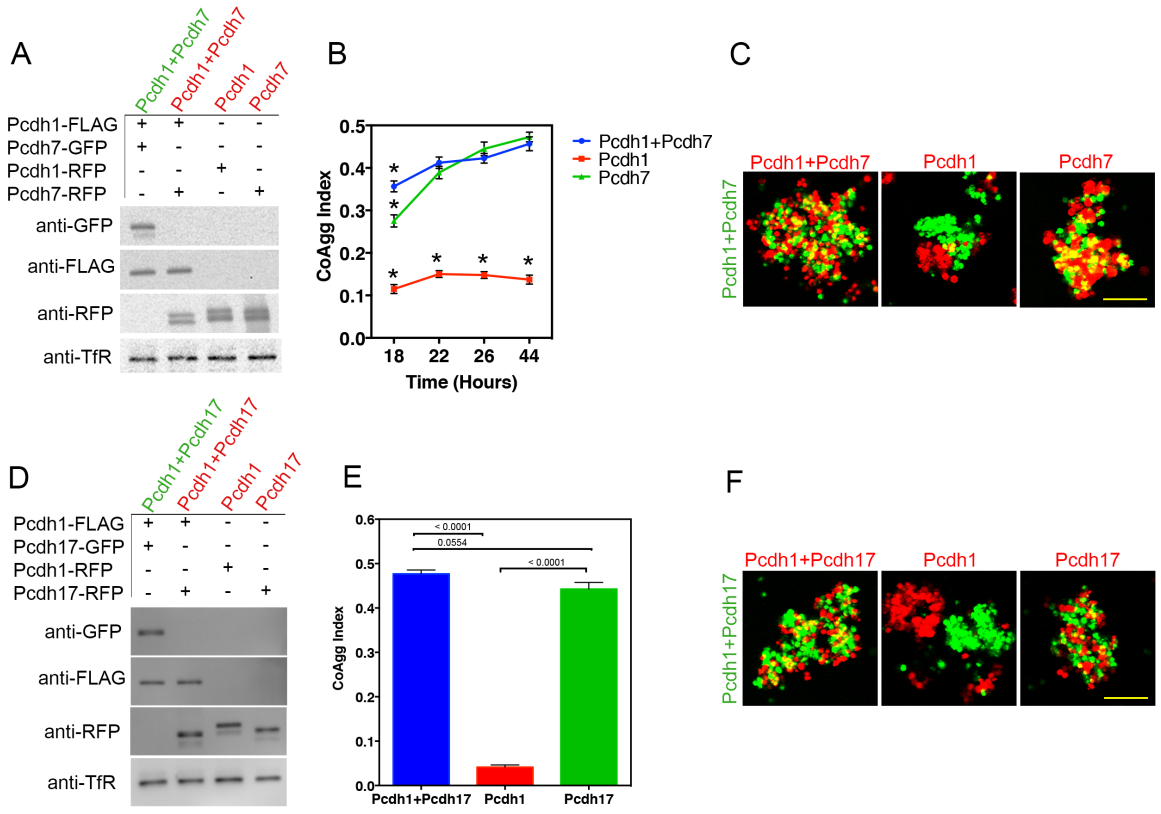

Figure 3.8 Deltas still maintain differential mixing behaviors in mismatched pairs even when variation in surface level expression is reduced. (**A**) Western blot of surface biotinylation for Pcdh1+Pcdh7 pairs showing that each gene is expressed at roughly equal levels at the cell surface. (**B**) CoAgg Index values for mismatched coaggregation assays for Pcdh1+7 cells mixed with Pcdh1+7, Pcdh1, and Pcdh7 cells across time points. Note that the Pcdh7 and Pcdh1+Pcdh7 cells display similar values while Pcdh1 cells are much lower. Differences between populations at the time points were highly significant by multiple t-tests (* = p-values < 0.0001). (**C**) Representative images of aggregates at 26 hours. Note that Pcdh1-RFP cells can only interface, while Pcdh7-RFP cells can intermix with Pcdh1+Pcdh7-GFP cells just as well as Pcdh1+Pcdh7-RFP cells. (**D**) Western blot of surface biotinylation for Pcdh1+Pcdh7 pairs showing that each gene is expressed at roughly equal levels. (**E**) CoAgg Index values at 26 hours, (F) Representative images, note that Pcdh1-RFP cells segregate from Pcdh1+Pcdh17-GFP cells whereas Pcdh17-RFP cells completely intermix. Each pair was tested across two independent electroporations. Error bars indicate standard error. P-values calculated using Student's t-tests. Scale bar=100 µm.

.

4.5. Higher Rotational Speeds Reveal Differential Apparent Affinities Among δ-Protocadherin

Our results suggest that deltas possess different apparent affinities because some are able to overcome the presence of a mismatched pair while others are not. For example, when Pcdh1+Pcdh7 are co-expressed, cells expressing Pcdh7 alone are still able to overcome the presence of Pcdh1 in the Pcdh1+Pcdh7 cells because the affinity of Pcdh7-Pcdh7 interactions is greater than Pcdh1-Pcdh1 interactions. To test this model a different way, we looked to prior studies by Steinberg, who demonstrated the importance of the rotational speed that cells are subjected to in cell aggregation experiments. When cells are incubated in round wells and subjected to gyratory movement, shear forces are imposed on the cells as they collide into one another. Using cell populations expressing different classical cadherins, he found that higher speeds increased the shear forces imposed on the cells, and therefore the stringency of coaggregation between different cadherin types. In contrast, extremely low speeds allowed weak hetero-cadherin binding to occur between two different cadherin populations that had previously not been observed (Duguay et al. 2003).

Unfortunately, most studies on protocadherin aggregation do not appear to take this factor into consideration, and either do not report the speed used (Thu et al. 2014) or only use high speeds (Emond et al. 2011; Cooper et al. 2016). As the selection of speed could drastically impact the results, it is important to fully consider the selection of speed used in aggregation experiments. For example, if a high speed is used that selects for

only the strongest bonds, weaker bonds that may still be biologically relevant will not be observed. Additionally, the *in vivo* relevance of aggregation at commonly used speeds such as 80RPM and greater remains to be explained. For these reasons, in all our experiments thus far we subjected the K562 cells to 15RPM, which is much lower than what is typically used. We also confirmed that our results with 15 RPM produced qualitatively equal results to cells that were not subjected to any rotational speeds (0 RPM), suggesting that 15 RPM does not create significant shear force to cause the loss of weaker bonds (data not shown). However, the slight agitation of 15 RPM allowed for the cells to better disperse throughout the well than at 0 RPM, enabling them to sample the population more efficiently and also enhance imaging quality.

We predicted that by subjecting δ-protocadherin aggregates to different rotational speeds, differences in apparent affinity among the deltas would be revealed. To test this prediction, we transfected three different populations of K562 cells (expressing Pcdh1, Pcdh7, or Pcdh17 alone), and ensured they expressed similar surface levels for each K562 population (Figure 3.9A). We subjected the cells to gradually increasing speeds (15, 120, 160, 200 and 220 RPM), and obtained confocal images for each time point. Here we assessed aggregate size (and not coaggregation), as each population was transfected with only a single delta. We predicted that the aggregates of deltas with lower apparent affinities would dissociate sooner and at lower speeds than deltas with higher apparent affinities. In preliminary studies, we determined that the precise order of events was important to obtain reproducible results. Different aggregates formed depending upon when the cells were subjected to higher speed (e.g. immediately after

transfection or after 24 hours), and how various speeds affected the size of aggregates formed by the different deltas (see methods).

At our standard 15 RPM aggregation speed, aggregate size for cells expressing Pcdh7 was consistently larger than that of Pcdh1 or Pcdh17. As we increased rotational speed, this relationship was maintained, and Pcdh7 aggregates remained larger than Pcdh1 or Pcdh17 aggregates (Figure 3.9B). Even at the highest rotational speed tested (220 RPM), small but distinct aggregates were still present in Pcdh7-expressing cells, while both Pcdh1 and Pcdh17 expressing cells had fully dissociated into single cells (Figure 3.9C).

Pcdh1, 7 and 17 were the top three of the mixing ranking from Figure 3.6C. We predicted that if we tested the lower ranking Pcdh19, it should have smaller aggregates at all speeds compared to the top three. However, because it was difficult to express Pcdh19 at the same high levels used in the Pcdh1, 7, and 17 experiment, we reduced Pcdh1 levels to match that of Pcdh19 (Figure 3.9D). We then carried out an identical experiment, and found that Pcdh19 always formed smaller aggregates than Pcdh1 at each speed (Figure 3.9E,F). This is consistent with the presence of a hierarchy of δ-protocadherin apparent affinities, with Pcdh7 having a higher apparent affinity than Pcdh1 or Pcdh17 (which were very similar), and with Pcdh19 having the lowest affinity among these four proteins. These rankings, generated using a different approach from those in Figure 3.6, still resulted in a similar hierarchy of delta affinities.

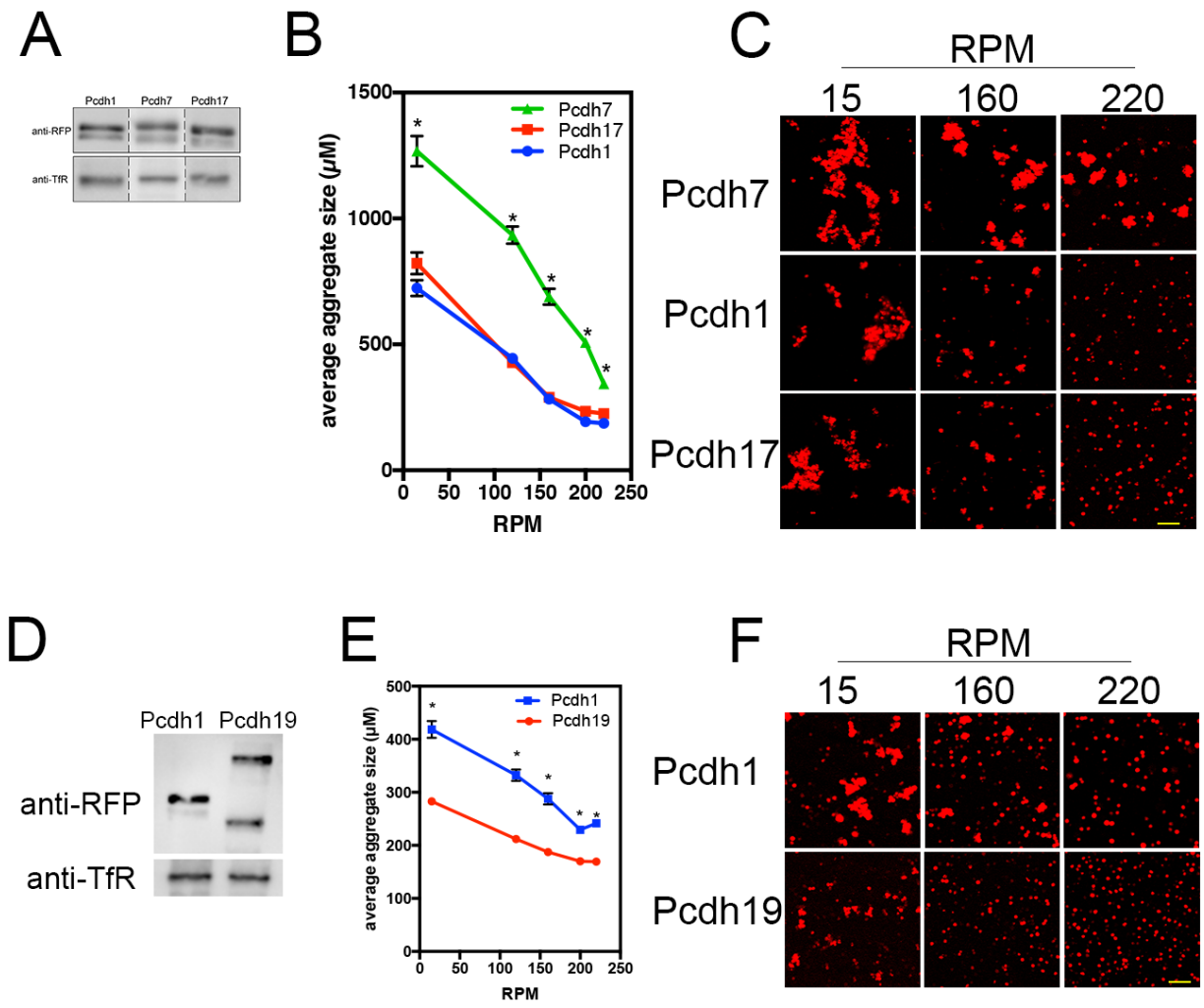

Figure 3.9 Increasing rotational speed on δ-protocadherin aggregates reveals differences in apparent affinities. (**A**) Western blot showing similar levels of surface protein levels for populations of cells expressing Pcdh1-RFP, Pcdh7-RFP, and Pcdh17-RFP. (**B**) Average size of aggregates as rotational speed is increased for Pcdh1, Pcdh7, and Pcdh17 cells. Note that Pcdh7 (green line) has consistently larger aggregates than Pcdh1 and Pcdh17 cells. P-values calculated by Student's t-test (* = $P > 0.0001$). (**C**) Western blot showing similar levels of surface protein levels for populations of cells expressing Pcdh1-RFP, and Pcdh19-RFP. (**E**) Average size of aggregates as rotational speed is increased for Pcdh1 and Pcdh19 cells. P-values calculated by Student's t-test (* = $P > 0.0001$). (**F**) Representative images at select time points. Note that Pcdh19 cells quickly dissociated into single cells compared to Pcdh1 cells. Each assay was repeated over four independent electroporations. Error bars indicate standard error. Scale bar = 100 μm.

.

We next assessed how rotational speed could be utilized to confirm differences in delta affinities observed with our mismatch coaggregation assay. In our initial mismatch experiment with Pcdh1 and Pcdh7, cells expressing Pcdh1 alone could only interface (index = 0.11) with cells expressing both Pcdh1+Pcdh7, whereas cells expressing only Pcdh7 intermixed highly (index = 0.33) with Pcdh1+Pcdh7 cells. We repeated the experiment but instead altered the rotational speed to see how the CoAgg Index values would change. Aggregates were allowed to form overnight at 15RPM and images were taken at 24 hours. The cells were then subjected to 120RPM for 1 hour and re-imaged (Figure 3.10A). The index of control experiments (Pcdh1+Pcdh7-GFP cells vs. Pcdh1+Pcdh7-RFP cells) dropped only slightly upon increasing speed but remained high (index = 0.29). In contrast, cells expressing Pcdh1 alone now failed to interface with the Pcdh1+Pcdh7 cells, and the two populations were now completely separated (index = 0.04). Finally, cells expressing Pcdh7 alone still intermixed with Pcdh1+Pcdh7 cells at the higher speed, in a manner highly similar to the control cells (index = 0.30) (Figure 3.10B). These results suggest that the interfacing behavior of Pcdh1 cells with Pcdh1+Pcdh7 cells is much weaker that the intermixing behavior of Pcdh7 cells with Pcdh1+Pcdh7 cells, enabling Pcdh7 cells to maintain adhesive interactions between mismatched populations even at higher speeds. These results are again consistent with Pcdh7 possessing a higher apparent adhesive affinity compared to Pcdh1.

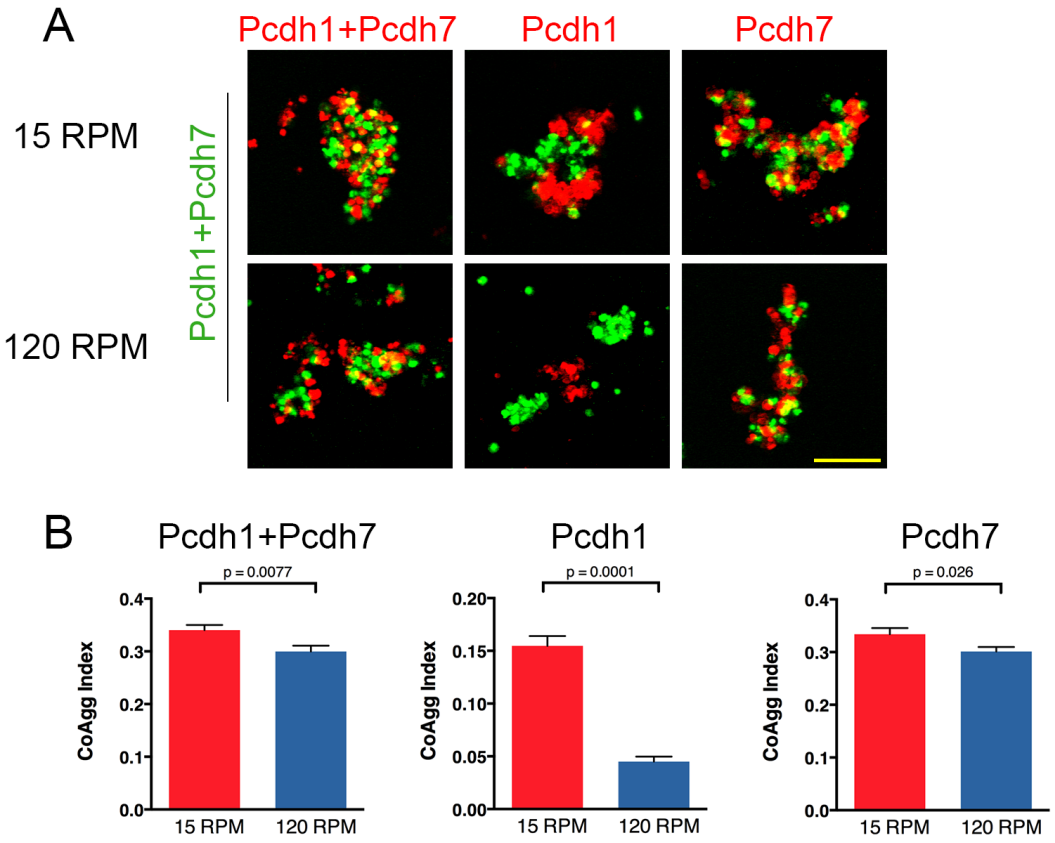

Figure 3.10 The impact of rotational speed on mismatch coaggregation assay with Pcdh1 and Pcdh7. At 15RPM (standard speed), Pcdh7-RFP cells intermix with Pcdh1+Pcdh-GFP cells equally well as Pcdh1+Pcdh7-RFP cells do, whereas Pcdh1-RFP cells will only interface. At increased rotational speed, the Pcdh1-RFP interfacing interactions are disrupted and the population now segregates, whereas Pcdh7-RFP cells remain intermixed, similar to the Pcdh1+Pcdh7-RFP control cells. (**A**) Representative images of mismatched coaggregation assays at both speeds, and (**B**) CoAgg Index values demonstrating change in coaggregation behavior between 15 and 120 RPM. P-values calculated with Student's t-test. Error bars indicate standard error. Scale bar = 100 μm.

4.6. δ-Protocadherin Extracellular Domains Have Low Sequence Identities

Our prior experiments showing differences in aggregate size, aggregate dissociation, and mismatched coaggregation among the deltas highly suggest they have different apparent affinities. To begin examining possible structural differences accounting for these observations, we annotated the signal peptide and transmembrane domains of each δ-protocadherin to isolate just the extracellular domains sequences, which are the regions that are responsible mediating adhesion. We performed MUSCLE alignments between all delta members and found they exhibit low sequence identity, around ~35% (Figure 3.11). Because delta-1s contain an extra EC domain, we then the compared sub-families (δ-1 vs δ-2). Sequence identity comparisons within sub-families exhibit slightly higher sequence identities but still remain fairly low (42-45%). A recent structural paper has shown that the adhesive *trans* interface of zebrafish Pcdh19 is mediated through a “forearm handshake” of EC domains 1-4 (Cooper et al. 2016). We therefore annotated each extracellular domain by utilizing the various common calcium binding motifs which are present between each domain as boundary markers, and compared domains 1-4 as well as each of the first four domains individually (Table 3.1). We found that overall, the first four domains also exhibit similarly low levels of sequence identities, but that the first domain is the lowest among all four domains. Together, this low level of sequence homology among the deltas is consistent with them possessing different apparent affinities.

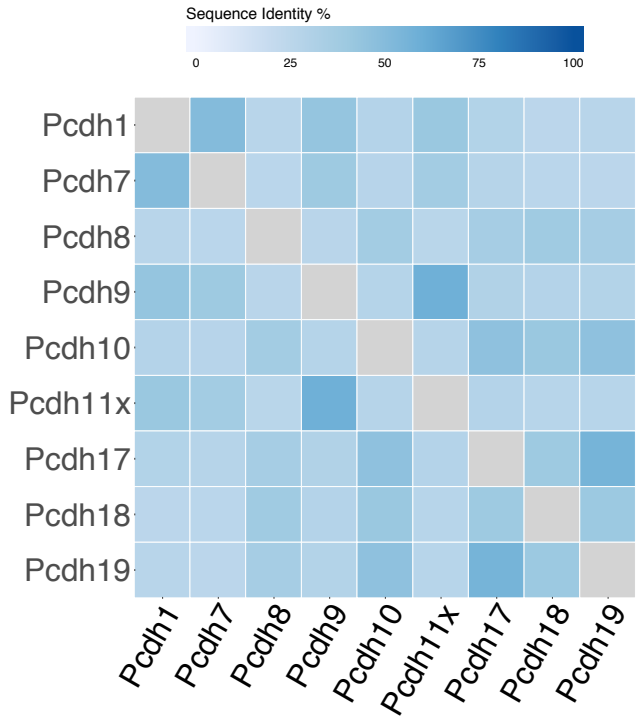

Figure 3.11 Pairwise protein sequence identities of δ-protocadherin extracellular domains. Note the low sequence identities.

Group	Full EC	EC 1-4	EC 1	EC 2	EC 3	EC 4
δ-1 (Pcdh1, 7, 9, 11x)	45.3	43.6	35.7	44.9	47.7	45.6
δ-2 (Pcdh8, 10, 17, 18, 19)	42.2	44.0	34.1	47.7	54.0	40.0
All deltas (δ-1 + δ-2)	34.4	37.5	29.1	38.7	44.4	37.4

Table 3.1 Pairwise protein sequence identity (%) of EC domains.

4.7. EC Domains 1-4 Mediate *trans* Binding of δ-Protocadherins

To further test the EC1-4 adhesive interface, we first created EC1-4 deletion constructs of Pcdh7 to confirm that this region mediates adhesion. Alone, these constructs failed to mediate any aggregation (Figure 3.12A), consistent with EC1-4 being the adhesive interface. We then repeated the mismatch coaggregation assay, but this time with cells co-expressing Pcdh7$^{(delEC1-4)}$ and Pcdh1 (Figure 3.12B). This deletion caused Pcdh7 cells to switch from intermixing to complete segregation when mixed with cells expressing Pcdh7$^{(delEC1-4)}$ (index = 0.01), consistent with the disruption of Pcdh7-Pcdh7 mediated adhesion. In contrast, Pcdh1 cells, which previously only interfaced with Pcdh1+Pcdh7 cells, could now intermix (index = 0.28) (Figure 3.12C). Together, this suggests that the EC1-4 plays a necessary role in δ-protocadherin adhesion. Additionally, it suggests that EC1-4 in Pcdh7 plays a dual role in the Pcdh1+Pcdh7 population, as it required to allow Pcdh7 cells to intermix, but also apparently acts to interfere with and prevent Pcdh1 cells from intermixing.

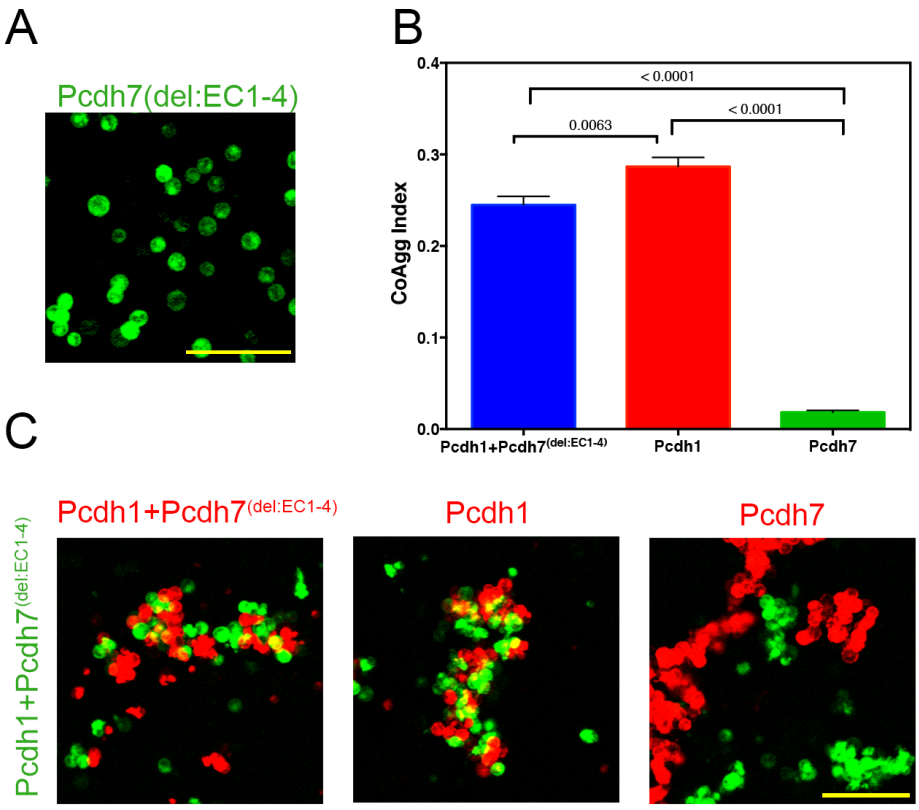

Figure 3.12 Extracellular domains 1-4 are required for *trans* binding of δ-protocadherins. (**A**) Deletion of EC1-4 in Pcdh7 removes adhesive function and aggregation formation. (**B**) CoAgg Index values of mismatch coaggregation assay using cells coexpressing Pcdh1 with Pcdh7$^{(del:1-4)}$. Pcdh7 cells, which usually intermix with Pcdh1+Pcdh7 cells, now segregate while Pcdh1 cells, which usually segregated, now intermix. (**C**) Representative images of coaggregation results. Each pair was tested over three independent electroporations. P-values calculated with Student's t-test. Scale bars = 100 µm.

To confirm whether the domains for adhesive specificity and affinity reside within EC domains 1-4, we created a chimera construct by swapping EC1-4 of Pcdh1 into Pcdh7 (Pcdh7$^{(EC1-4:Pcdh1)}$). We first addressed specificity by testing aggregate formation between cells expressing the Pcdh7$^{(EC1-4:Pcdh1)}$ and cells expressing Pcdh1 or Pcdh7 alone. We found that Pcdh7$^{(EC1-4:Pcdh1)}$ now completely intermixed with Pcdh1, and

segregated from Pcdh7, confirming that residues for adhesive specificity are within EC1-4 (Figure 3.13A,B). We then expanded this to our mismatch coaggregation assay. As expected, cells that co-express Pcdh1+Pcdh7$^{(EC1\text{-}4:Pcdh1)}$ now intermixed with Pcdh1 cells, but segregated from Pcdh7 cells. These results suggest that EC domains 1-4 contain the residues responsible for δ-protocadherin specificities and apparent affinities.

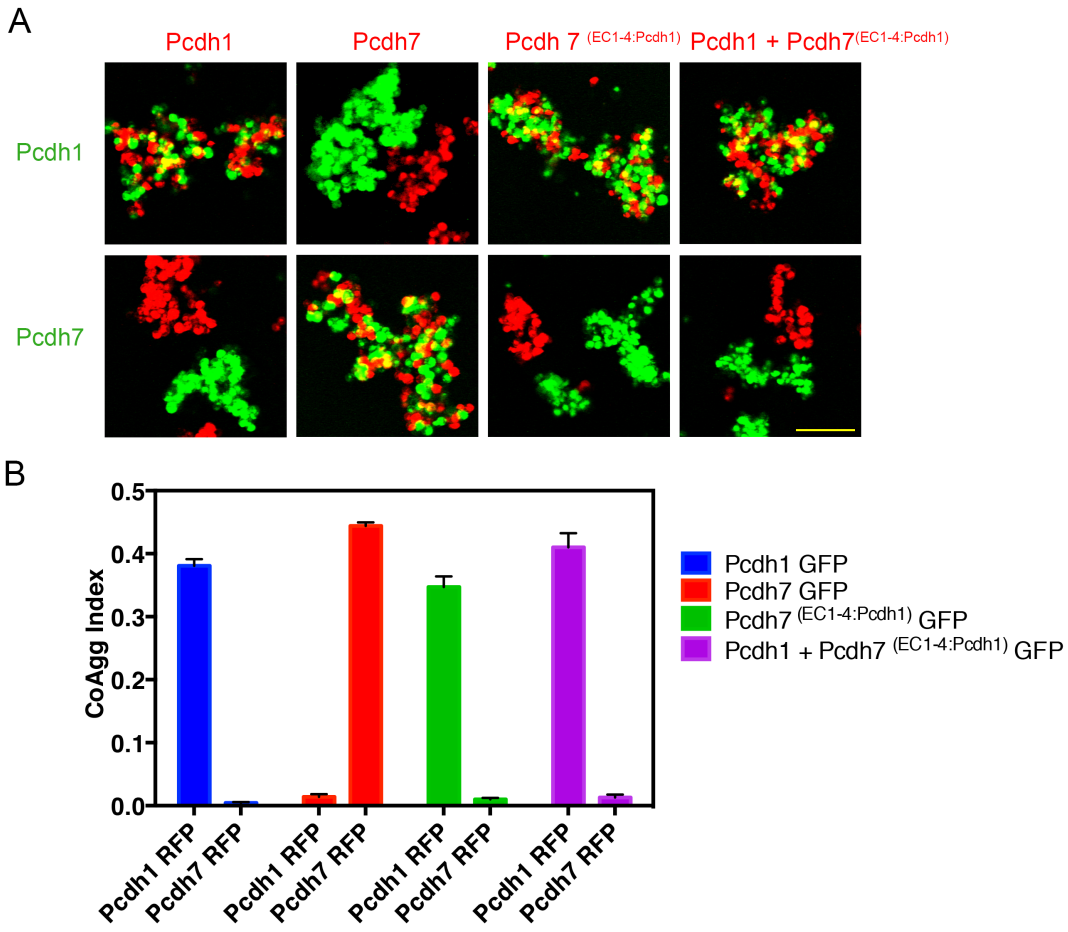

Figure 3.13 Extracellular domains 1-4 confer adhesive specificity. (**A**) Swapping Pcdh1 EC1-4 into Pcdh7 is sufficient to switch *trans* binding specificity, as well as coaggregation behavior in mismatch coaggregation assay. (**B**) Coaggregation Index values for each pair tested. Note the dramatic reversal when EC1-4 is swapped. Each pair was tested over two independent electroporations. Error bars indicate standard error. Swaps scale bar = 100 µm.

4.8. Modulating Surface Level Expression of δ-Protocadherins Alters Aggregation Behaviors

Our results above highly suggest that the members of the δ-protocadherins have different apparent affinities. Most of the experiments performed to support this idea were controlled in a manner such that differences in surface expression level were reduced and each protocadherin had an equal opportunity to influence the outcome. However, our *in situ* and single cell expression analyses (see Chapter 2) suggest that individual neurons express δ-protocadherins at various levels. We therefore explored how modulating expression might influence the apparent adhesive affinity differences among the deltas when expressed alone, and in combination.

We first optimized expression levels for Pcdh1, 7, and 17 alone to generate similar gradients of "low," "medium," and "high" surface levels of expression (Figure 3.14A). To determine if apparent adhesive affinity is dependent upon surface level, we electroporated K562 cells expressing low, medium or high level of Pcdh7-RFP and Pcdh17-RFP separately, and subjected them to the same speed assay as in Figure 3.9 Analysis of the aggregate sizes for these experiments showed that when comparing each expression level, Pcdh7 always had larger aggregate sizes at every speed (Figure 3.14B-C). Even Pcdh7 at medium levels maintained larger aggregate sizes than Pcdh17 at high levels. These patterns are consistent with Pcdh7 having a greater apparent affinity than Pcdh17.

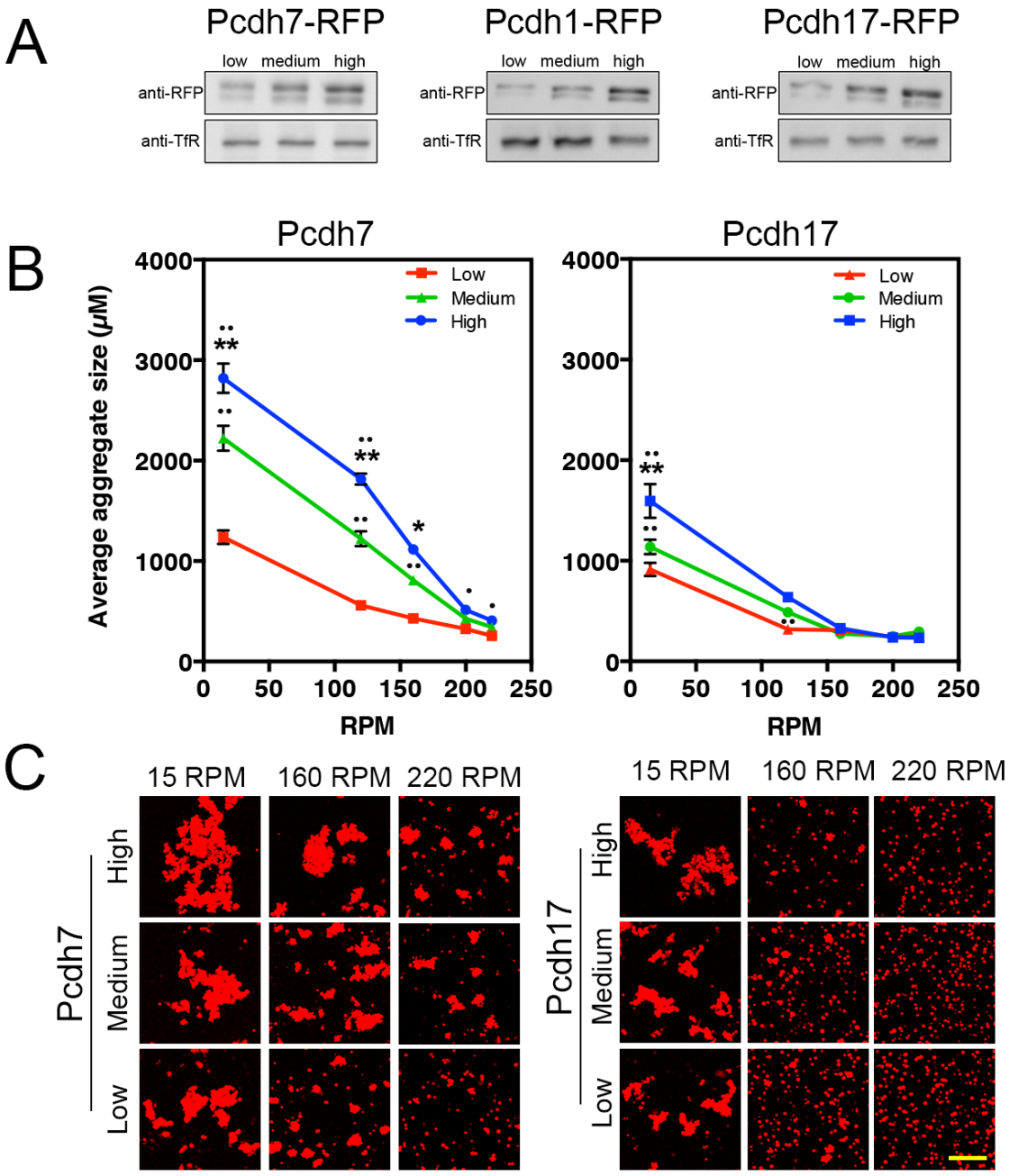

Figure 3.14 (**A**) Western blots showing establishment of gradients (high, medium, and low) of surface membrane levels for Pcdh1, Pcdh7 and Pcdh17. (**B**) Average aggregate size of cell populations expressing gradients of Pcdh7 and Pcdh17 at increasing rotational speed. P-values calculated by multiple Student's t-test (** = statistically significant from medium, $p > 0.0001$, * = $p > 0.005$, •• = statistically significant from low, $p > 0.0001$, • = $p > 0.05$). (**C**) Representative images of aggregates at select speeds. Each assay was performed over four independent electroporations. Error bars indicate standard error. Scale bar = 200 μm.

.

We next transfected K562 cells with low, medium, and high levels of Pcdh1, 7, and 17 and performed aggregation assays where the type of protocadherin was matched between populations, but the levels of that protocadherin varied. For example, Pcdh1-GFP (medium) was mixed with cells expressing Pcdh1-RFP at low, medium and high levels. This experiment was repeated for Pcdh7 and Pcdh 17 (Figure 3.15A,B). Notably, cells expressing matched levels of a given delta (e.g. both populations expressing medium levels of protein) had consistently higher intermix values than when the two populations had differing, non-matched levels of surface expression (e.g. medium vs. low or medium vs. high). Additionally, we consistently identified higher index values for Pcdh7 relative to Pcdh17 and Pcdh1 regardless of the level of expression of each gene, consistent with Pcdh7 having the highest apparent affinity among these deltas.

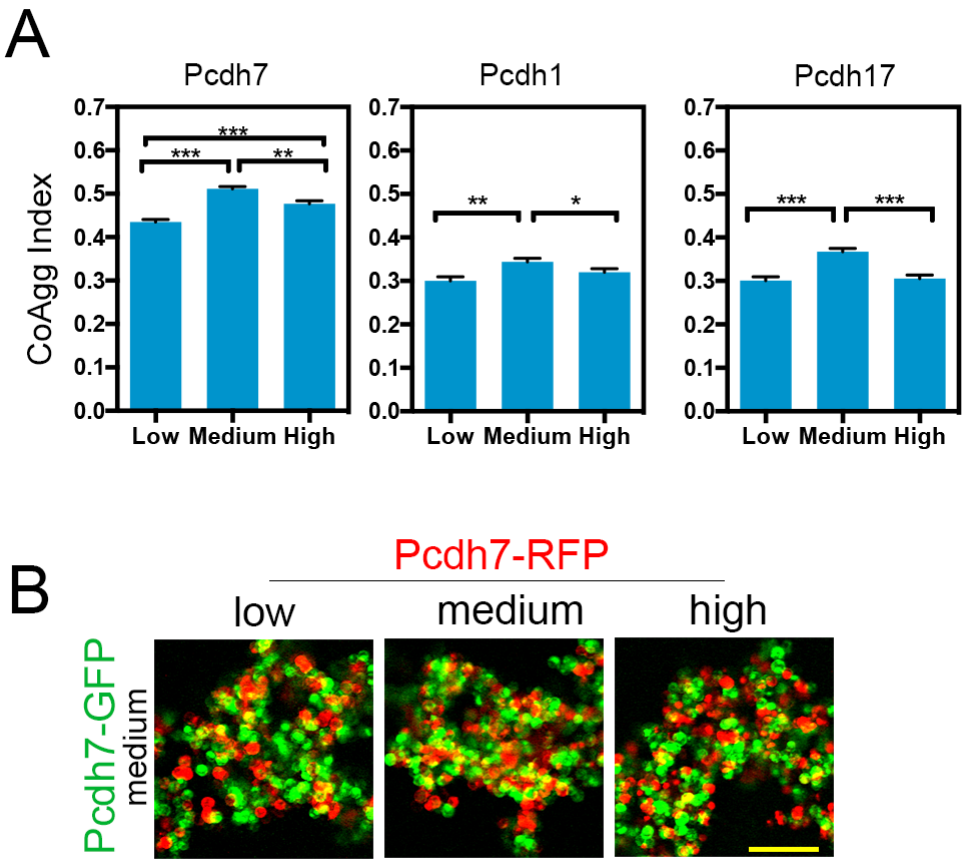

Figure 3.15 Populations of matched deltas show slight preference towards matched levels. (**A**) CoAgg Index values for medium populations allowed to mixed with low, medium or high levels populations. In each case, the medium-medium pair had the highest CoAgg Index. (**B**) Representative images of Pcdh7-GFP medium cells coaggregated with Pcdh7-RFP cells at low, medium and high levels. Subtle differences can be seen when comparing the results of the medium cells to the low and high cells. P-values calculated with Student's t-test, *** = $p < 0.0001$, ** = $p < 0.001$, * = $p < 0.05$. Error bars indicate standard error. Scale bar = 100 μm.

4.9. Manipulation of Surface Levels in Populations Expressing a Single δ-Protocadherin Does Not Change Coaggregation Behavior in the Mismatch Coaggregation Assay

Our results so far suggest that deltas possess different relative affinities, but that the amount at which they are expressed can influence aggregate size and slightly alter homophilic coaggregation. We next asked how varying surface levels would influence a mismatch coaggregation assay. We had previously shown that when differences in expression level are minimized, mixing Pcdh1+Pcdh7-GFP cells with Pcdh1-RFP cells led to interfacing while Pcdh7-RFP cells highly intermixed. In a second experiment, Pcdh1+Pcdh17-GFP intermixed with Pcdh17-RFP cells, while segregating from Pcdh1-RFP cells (see Figure 3.8). We used the same gradients used to establish low, medium and high levels in Figure 3.15 to test whether the index values would be affected by changes in the surface level expression of the RFP populations that express a single protocadherin. This experiment would test, in a mismatch coaggregation assay, whether or not the expression level on cells expressing a single delta would influence its ability to interact with cells expressing two deltas at equal levels. For example, could a population expressing a given delta that interfaced with a mismatch population at 'balanced' levels intermix if it expressed a lot more? Although some changes in index values were observed, they were overall remarkably stable. Importantly, the CoAgg index category never changed. For example, although the cells expressing low levels of Pcdh7-RFP had reduced CoAgg index values, compared to medium and high, each population still intermixed (Figure 3.16 A,B). Consistent with this, increasing the surface

level of Pcdh1-RFP cells increased the CoAgg Index value, but it always remained interfacing (Figure 3.16C,D). Further tests with Pcdh1+Pcdh17 yielded results consistent with that of the Pcdh1+Pcdh7 (Figure 3.16E-H).

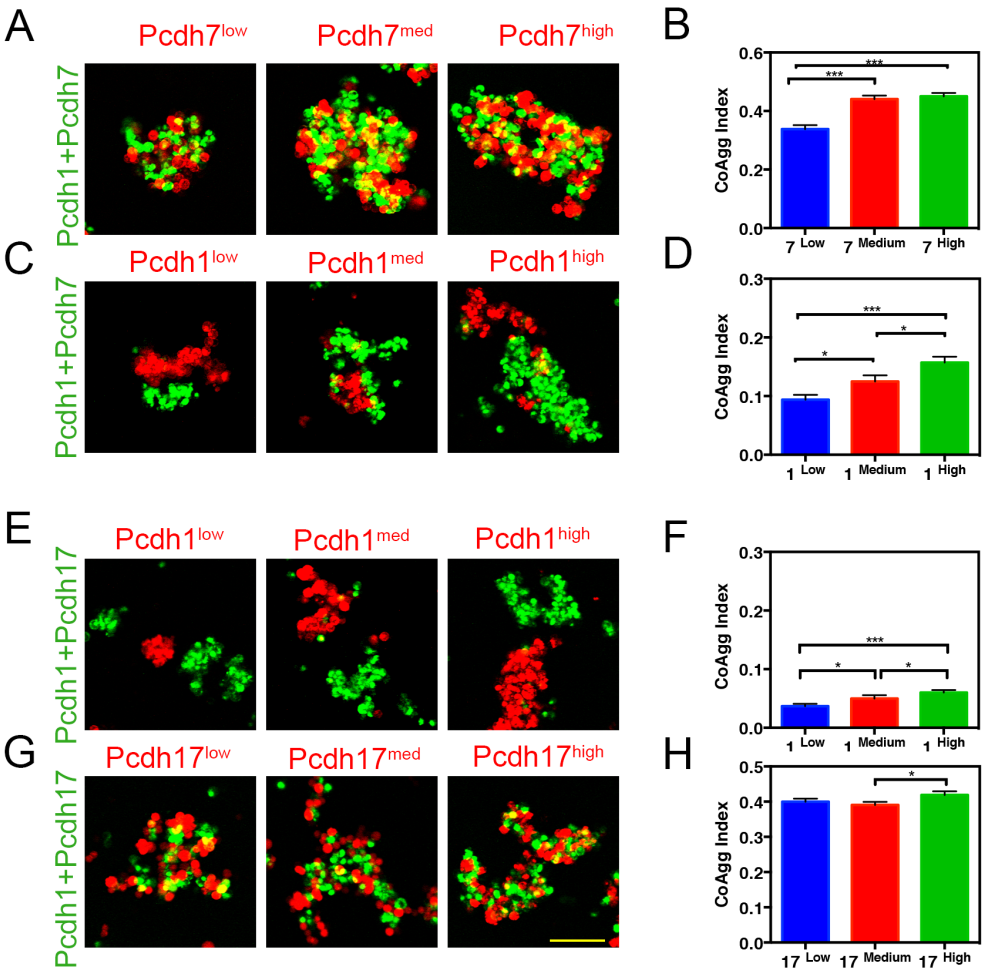

Figure 3.16 Altering level of a delta in the population expressing a single delta in the mismatch coaggregation assay does not significantly alter outcome of coaggregation behavior. (**A**) Cells expressing Pcdh7-RFP at low, medium and high surface levels consistently mix with Pcdh1+Pcdh7-GFP cells. (**B**) Although CoAgg Index values changed slightly, each population still intermixed. Gradients tested for the single pair did not change coaggregation behavior for (**C,D**) Pcdh1 vs Pcdh1+Pcdh7, (**E,F**) Pcdh1 vs Pcdh1+Pch17, and (**G,H**) Pcdh17 vs Pcdh1+Pcdh17. Each pair was tested across two independent electroporations. P-values determined by Student's t-test, *** = $p < 0.0001$, ** = $p < 0.001$, * = $p < 0.05$. Error bars indicate standard error. Scale bar = 100 µm.

.

4.10. Manipulation of the Relative Proportion of Surface Expression in Populations Expressing Two δ-Protocadherins Changes Coaggregation Behavior in Mismatch Coaggregation Assay

What is the relationship between the number of deltas expressed within a neuron, their apparent adhesive affinities, and their relative expression to one another? In Fig. 3.8, we performed mismatch coaggregation experiments while maintaining approximately equal levels of all deltas within all cells. In contrast, in Figure 3.16, we maintained approximately equal levels of deltas expressed within cells expressing two δ-protocadherins, but varied the level of delta expression in the single population. We next maintained a constant level of delta surface expression in the single population, but varied the proportion of surface level expression in the double population to ask what effect varying this ratio could have on the adhesive interactions between two populations of cells.

To answer this question, we tested nine pairs of δ-protocadherins. To perform these experiments, we generated six different populations for each pair tested. For example, to test the effect of altering the proportion of Pcdh1 and Pcdh7 when co-expressed, we electroporated two different populations of Pcdh1+Pcdh7. In the first population, the amount of DNA input of Pcdh1-GFP was much higher than that of Pcdh7, and so we called this population Pcdh1^{high}+Pcdh7^{low}-GFP. In the second population, the DNA input was flipped such that more Pcdh7 was used than Pcdh1 (Pcdh1^{low}+Pcdh7^{high}-GFP). To test for coaggregation with these two GFP populations, we also generated identical RFP

populations for each pair, as well as populations expressing Pcdh1-RFP and Pcdh7-RFP individually at a medium level in between the high and low expression levels. For each pair tested, every GFP population was then tested against every RFP population, and their index values were obtained. These experiments generated two components within each pair. First, it tests how surface expression proportions affect mixing when the genes were matched (i.e. double expressing GFP populations vs. double expressing RFP populations). Secondly, it tests the influence of proportion in the double cells in the mismatch coaggregation assays (i.e. double expressing GFP populations vs. single RFP populations).

As expected, control pairs that had matched expression levels (i.e. Pcdh1^{low}+Pcdh7^{high}-GFP vs. Pcdh1^{low}+Pcdh7^{high}-RFP) intermixed and had high CoAgg index values (Figure 3.17A first four squares, Figure 3.17B). Differences in the CoAgg Index values between the two sets of control pairs could often be observed, depending on which genes were being tested. Importantly, populations with matched δ-protocadherins but unmatched levels had significantly reduced index values. For example, when levels between the populations were matched for Pcdh7+Pcdh17, the index values were measured to be 0.36, but 0.05 when unmatched. This indicates that matched surface expression levels can be just as critical as matching the type of deltas.

For the single-RFP populations, their outcomes were significantly dependent on which population of the co-expressed GFP pair they were being tested against. For example, Pcdh7^{high}+Pcdh1^{low} cells intermixed with Pcdh7-RFP cells but segregated from Pcdh1-RFP medium cells. In contrast, Pcdh7^{low}+Pcdh17^{high} cells segregated from Pcdh7-RFP

cells and intermixed with Pcdh17-RFP cells (Figure 3.17, last four squares). Thus, the proportion of the GFP population was sufficient to reverse the outcome, even though both genes were being expressed. For example, Pcdh7^{high}+Pcdh1^{low} cells intermixed with Pcdh7-RFP cells and segregated from Pcdh1-RFP cells (Figure 3.17A, top row). In contrast, Pcdh7^{low}+Pcdh1^{high} cells interfaced with Pcdh7-RFP cells and while intermixing with Pcdh1-RFP cells (Figure 3.17A, bottom row). Thus, the proportion of deltas in the GFP population was sufficient to alter the coaggregation outcome, even though both genes were being expressed. These results suggest that the relative amount of expression can critically modulate the coaggregation outcome, with the level and affinity of a given delta acting as a competitor to the other deltas expressed within the cell. Several other pairs of deltas were tested in this manner (Figure 3.17 C-F, Figure 3.18) and each experiment produced similar results. In theory, changing the relative proportion of co-expressed deltas could be one method for fine tuning adhesive interactions.

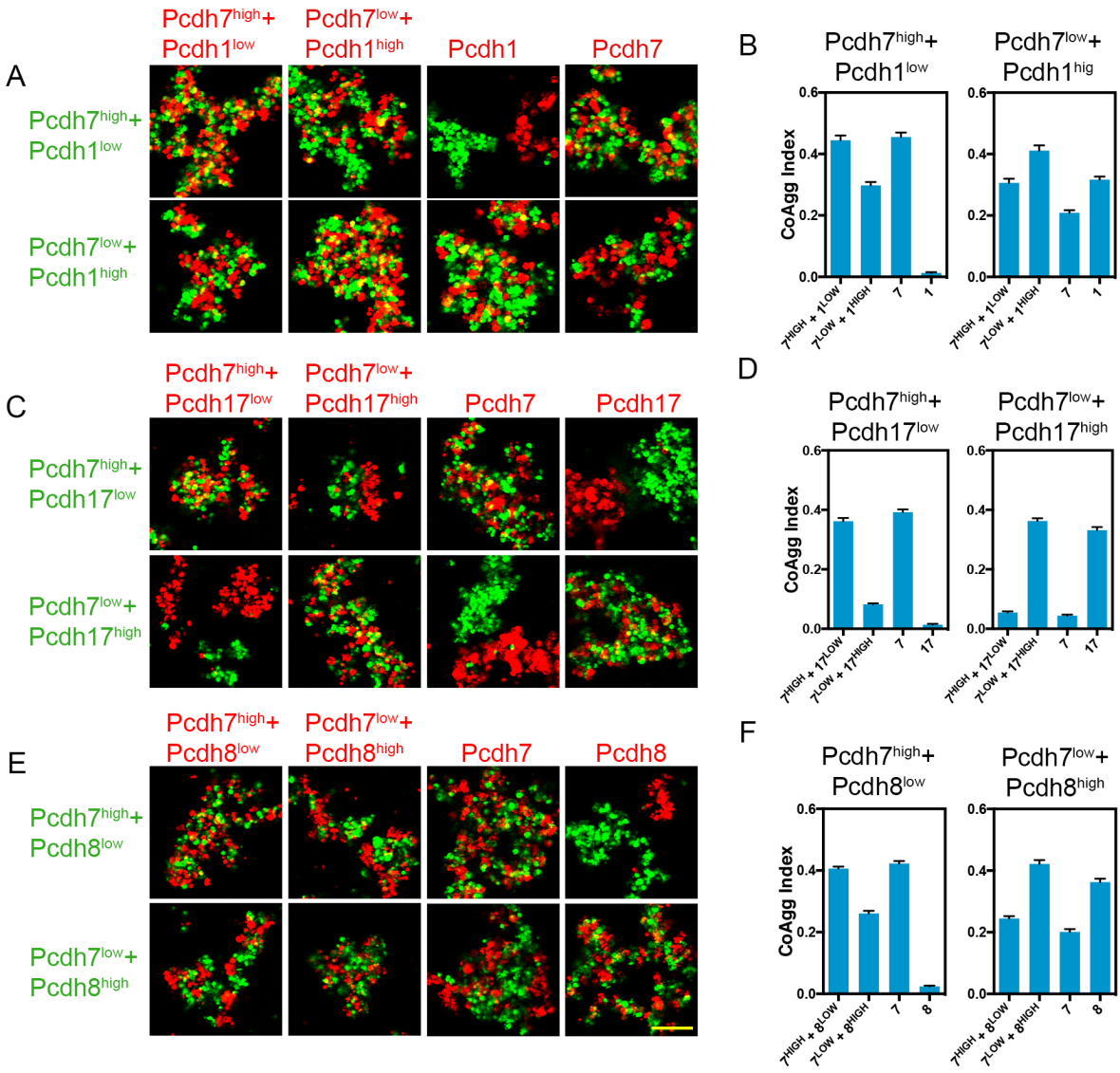

Figure 3.17 Changing the proportion of surface level in the population that expresses two deltas drastically changes coaggregation behavior with other cell populations. (**A**) Coaggregation behavior of Pcdh7^{high}+Pcdh1^{low} cells with cells expressing Pcdh7^{high}+Pcdh1^{low}, Pcdh1, or Pcdh7 (top row) changes when changed to Pcdh7^{low}+Pcdh^{high} (bottom row). (**B**) Coaggregation Index values of each RFP population when allowed to mix with each GFP population. Note that each pair changes depending on which high/low pair it is being mixed with. Similar effects were seen with (**C,D**) Pcdh7+Pcdh17 and (**E,F**) Pcdh7+Pcdh8. Each pair was tested across two independent electroporations. Error bars indicate standard error. Scale bar = 100 µm.

.

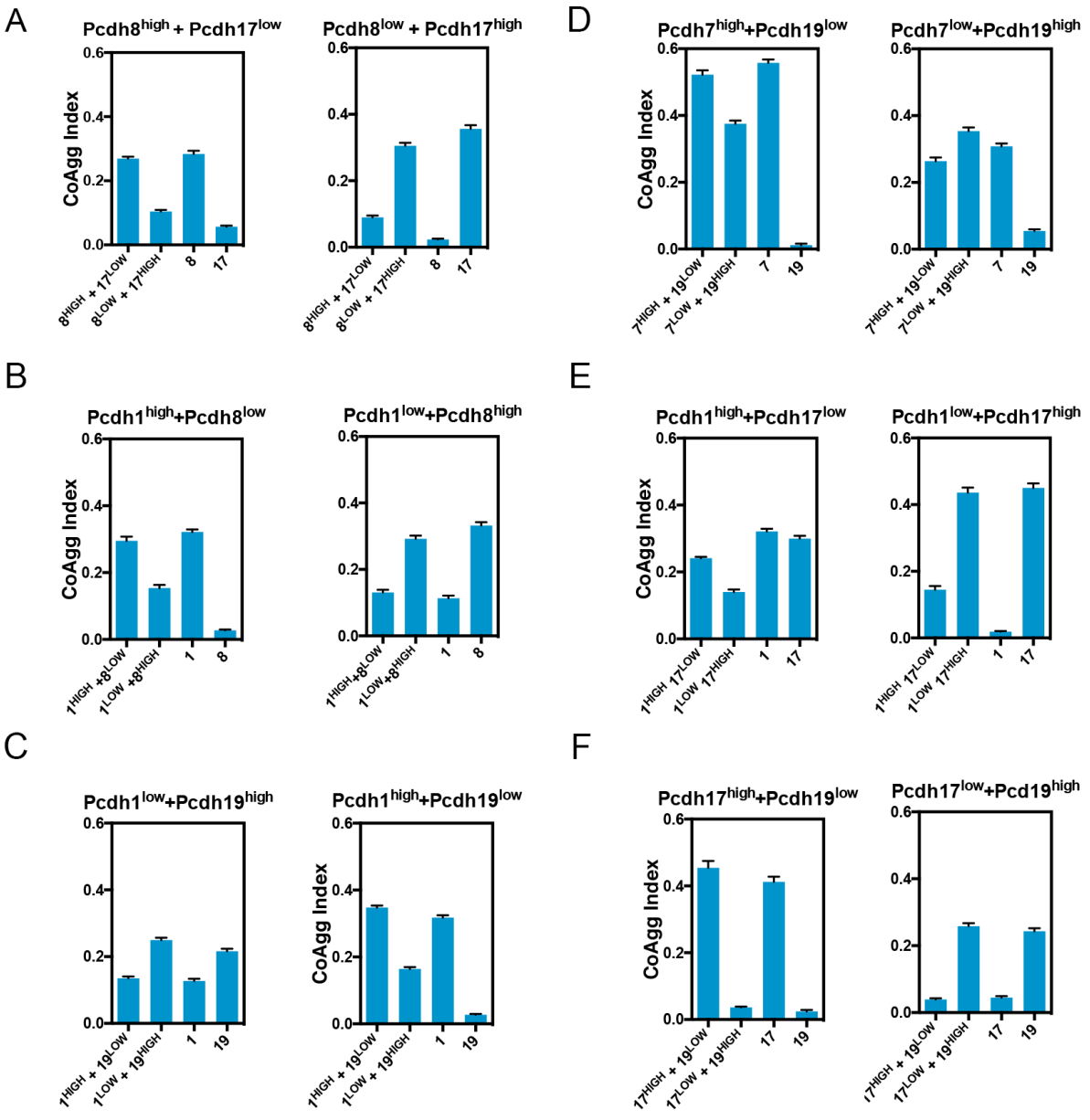

Figure 3.18 Coaggregation Index values of six more pairs of deltas testing impact of switching high/low status in cells expressing two deltas. Note that the index value for a cell population changes depending on which high/low pair it is mixed with. (**A**) Pcdh8+Pcdh17, (**B**) Pcdh1+Pcdh8, (**C**) Pcdh1+Pcdh19, (**D**) Pcdh7+Pcdh19, (**E**) Pcdh1+Pcdh17, and (**F**) Pcdh17+Pcdh19. Each pair was tested across two independent electroporations. Error bars indicate standard error.

Finally, to confirm that the differences in input DNA altered the proportion of the surface expression in these experiments, we checked a select number of pairs by performing surface biotinyation and western blots. We then approximated the differences between the high/low pairs by performing band analysis to produce a ratio of GFP to RFP signal (Figure 3.19) In most cases, the measured surface levels corresponded to the “low” and “high” classifications of the DNA input. However, in some cases (e.g. Pcdh7-GFP+Pcdh19-RFP) the pairs do not flip, but the proportion still changes. These results are consistent with the cellular behavior observed in Figures 3.17 and 3.18 and support the idea that the relative proportion of each delta, when co-expressed with others, is a critical component in modulating their adhesive specificities with other populations.

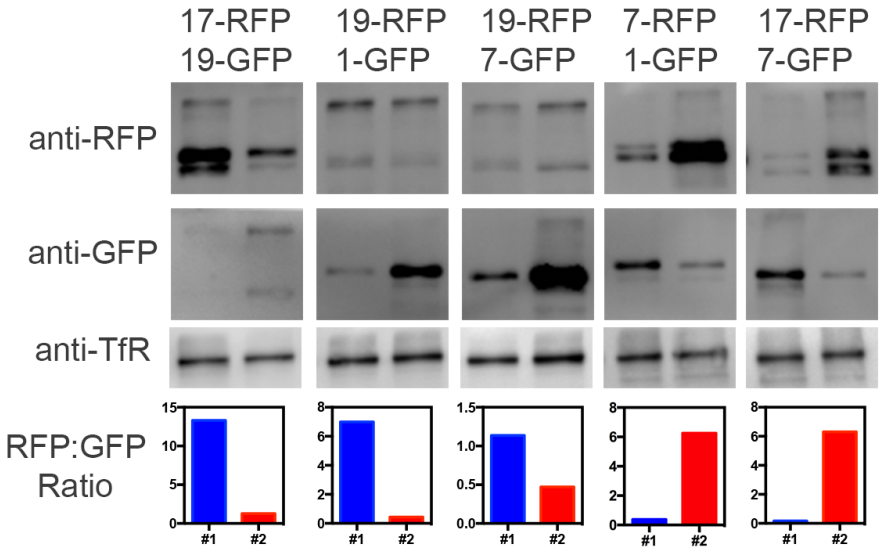

Figure 3.19 Labeling and detection of δ-protocadherins at the cell surface for select pairs from Figures 3.17 and 3.18 demonstrating change in surface level proportions between high/low DNA input. Protein samples were concentrated to enhance detection of deltas being expressed at lower levels, and this increased the visibility of higher molecular weight bands.

5. Discussion

Previous studies have shown the important roles the δ-protocadherins play in neural development (Uemura et al. 2007; Dibbens et al. 2008; Emond et al. 2009; Williams et al. 2011). However, they have never before been studied as a family to provide a systematic understanding of how they might function together. In Chapter 2, we showed that randomly selected individual OSNs express a wide range of δ-protocadherin combinations, raising important questions about how different combinations will influence their adhesive functions with other neural populations. For example, how might combinatorial expression be useful for a neuron? What logic may underlie the function of the combinations expressed, if any? And does the precise identity of each combination matter or are there functional redundancies? In this chapter, we systematically examined δ-protocadherins individually, and in simple combinations to begin to addressing these questions.

In stark contrast to the clustered protocadherins, which have been shown to always mediate segregation between populations that are mismatched by just one protocadherin, we show that delta-protocadherins are much more diverse in their possible outcomes when they begin to be expressed in simple combinations. Importantly, we report that the adhesive behaviors of populations expressing multiple deltas depends highly on which deltas are being expressed. Through several experiments, we propose that this is due to differences in apparent affinity among the deltas, but that the relative expression level of each delta modulates the influence of

these affinity differences. Together, we propose a working model which posits that δ-protocadherin combinatorial adhesive codes are generated not only through combinatorial expression, but also the relative affinity and expression levels of the specific delta repertoire. Our results supporting this model are discussed below.

5.1. δ-Protocadherins Mediate Specific *trans* Homophilic Adhesion

The δ-protocadherins have often been proposed to mediate combinatorial adhesive codes for neural development (Wöhrn et al. 1999; Hertel et al. 2008; Krishna-K et al. 2011), yet the adhesive characteristics of each delta had never been systematically studied. Using the K562 cell aggregation assay, we characterized each delta individually and then asked if populations of deltas only engage in aggregation with cells expressing the same protocadherin, or if they could cross-adhere with different deltas on different cells. By testing all possible pairs, we concluded that each delta is strictly homophilic in *trans*. This suggests that δ-protocadherins cannot stably "cross-adhere" and that any adhesion mediated by a delta on one cell must be interacting with the same type of delta on an opposing cell. Thus, the fundamental adhesive unit is the homophilic *trans* interface. Together, these adhesive characteristics are entirely consistent with the work that has been done on the clustered protocadherins.

5.2. δ-Protocadherins Possess Different *trans* Apparent Affinities

5.2.1. Evidence from Differential Coaggregation Behaviors in Mismatched Populations

Since we showed that deltas are expressed in combination in single neurons (see Chapter 2), and many have proposed that this could generate adhesive codes, we began to address how their adhesive functions may change when more than one delta are expressed per cell. In prior studies on the clustered protocadherins, any mismatch that existed between two cell populations would cause them to form distinct, segregated aggregates (Schreiner and Weiner 2010; Thu et al. 2014). This was tested and found to be true when even up to five were expressed. Our systematic screen of mismatch coaggregation assays was intended to test simple combinations and see if the deltas behave similar to the clustereds in this respect. In these experiments, two deltas were co-expressed in one population, and the ability of different populations expressing just one delta from the pair to coaggregate with the population co-expressing two deltas was observed. In striking contrast to the clustered protocadherins, we unexpectedly found that the δ-protocadherins produced diverse outcomes - some pairs of populations intermixed completely, some interfaced, while others segregated. Importantly, the outcome depended on which deltas were being tested. To quantify and order these variable outcomes, we developed and utilized a novel coaggregation measure, the Coaggregation Index. To control for the possibility that our diverse results were not the result of skewed or severely unbalanced co-expression in the cell populations that were electroporated with two protocadherins, we titrated and measured the cell surface

expression of select pairs to ensure that expression levels were roughly equivalent. After reducing the possibility that variation of cell surface expression was the cause of our results, we still observed differences among coaggregation behaviors, highly suggesting that different δ-protocadherins possess different apparent affinities.

5.2.2. Evidence from Differential Response to Increased Rotational Speeds

Although it is known that rotational speed can play a critical factor in cell aggregation experiments, many studies do not address the role of rotational speed in the result of aggregation experiments. At higher speeds, shear forces are increased, and weaker bonds begin to dissociate. We decided to use this overlooked principle as a different way to confirm the differences in apparent affinities of the deltas. We tested higher rotational speeds with deltas alone, and also with the mismatch coaggregation assay. In all cases tested, our results were consistent with our initial hierarchical rankings - deltas proposed to have weaker apparent affinities dissociated more easily at higher speeds that deltas proposed to have higher apparent affinities. These results further support the idea that the δ-protocadherin family members possess a range of apparent affinities that causes them to have differentiating cell adhesion functions.

5.2.3. Evidence from Sequence, Structure and Domain Swaps

Proteins that exhibit different apparent affinities should be able to be explained by their underlying structural differences. Our sequence analyses of the extracellular domain regions of the deltas showed low sequence identity (less than 40%). Upon examining the sequences more closely, we observed that some of the higher affinity ranking deltas

such and Pcdh7 and Pcdh17 had noticeable regions where the sequence was extended and not present in others. These extensions are found within the first four domains, consistent with recent structural studies showing that the clustered protocadherins form *trans* binding through EC 1-4 (Rubinstein et al. 2015; Goodman et al. 2016b; Nicoludis et al. 2016; Goodman et al. 2016a; Cooper et al. 2016). We therefore performed select deletion and swap constructs, and proved that domains controlling for specificity and affinity are contained within EC1-4, consistent with these prior structural studies. Together, these analyses further support our data by showing that there are sufficient differences in their sequences to mediate functional difference in their apparent affinities, and that the critical residues likely reside within EC domains 1-4. In contrast, the clustered protocadherins display higher extracellular sequence identity (~60%) and do not display any gaps or insertions in their alignments. Thus, it is entirely possible that the sequence differences among the deltas create a hierarchy of affinities. In contrast, the clustereds are all more similar in their structures and even though they display high specificities, may not exhibit such a wide range of different affinities.

Taken together, our aggregation experiments highly suggest that the differences in the behavior of deltas we observed are the result of deltas possessing different apparent affinities. Therefore, which deltas a neuron expresses becomes important because they are not all equal in their adhesive affinities. Future biochemical and structural studies will be required to actually define the affinities of these adhesive interfaces and confirm the extent of their differences.

5.3. Modulation of Affinity Repertoire by Surface Expression Level

If deltas have different affinities, which are fixed, what is the utility of expressing them in combination? For example, if Pcdh7 has the highest affinity such that it will always outweigh weaker affinities, even when combined, what use would that provide for neurons? Such a system would defeat the purpose of combinatorial expression. We showed in Chapter 1 through single-cell analysis and *in situ* hybridization that OSNs can express deltas at various levels. We therefore examine the effect of varying the surface level of protocadherins to see how the outcome of our coaggregation experiments might change. We concluded that the relative surface level between co-expressed deltas is a critical way to modulate their differences in apparent affinities.

5.3.1. Effect of Surface Level on Homophilic Populations

Studies on classical cadherins have shown that cell populations expressing the same cadherin but at different levels can sort out, thus showing the importance of surface level (Duguay et al. 2003). To date, the impact of surface level has never been reported for any protocadherin. By establishing surface expression gradients of low, medium, and high for Pcdh1, Pcdh7, and Pcdh17, we compared the CoAgg Indices of medium-low and medium-high to medium-medium pairs for each delta. Under our experimental conditions, these level differences did not cause segregation or interfacing and the populations still intermixed. However, we did notice that unmatched levels displayed reduced intermixing compared to the matched medium-medium pair. Although the effects are relatively subtle, they are statistically significant and consistent with the idea

behind Steinberg's Differential Adhesion Hypothesis (DAH) that cell populations will try to reach a state of minimized interfacial free energy. It can also be interpreted from a simple probabilistic stance - cells expressing a certain level of molecules of a delta are more likely to come into contact and adhere with cells expressing the same level than one expressing fewer molecules. We also note that all pairs of Pcdh7 cells had higher index values than all the Pcdh1 and Pcdh17 pairs, and that Pcdh1 and Pcdh17 had similar values, again consistent with our initial hierarchical rankings of apparent affinity. By testing these gradients against increasing rotational speeds, we confirmed that more surface proteins led to bigger aggregates, but that affinity rankings remained constant at each gradient level. These results suggest that differences in level can effectively modulate the aggregation and coaggregation outcomes even when the populations express the same delta and thus the same overall apparent affinity.

5.3.2. Effect of Surface Levels in Mismatched Populations

How could surface expression influence coaggregation between mismatched populations? We took pairs that we had titrated to express roughly equal surface levels (e.g. Pcdh1+Pcdh7) and tested them in mismatch coaggregation assays with cells expressing different levels of a single delta. Somewhat unexpectedly, having more or less surface protein on the single population did not drastically change the outcome. Although slight differences were observed, the category (e.g. interfacing) of coaggregation never changed. For example, if a single population intermixed when the levels were all roughly equal with the co-expressed pair, it still intermixed even when it was expressed at a lower level. This suggests that the levels of the single population do

not affect mismatched coaggregation. In contrast, if the proportion of the co-expressed deltas are altered, dramatic changes are observed in coaggregation behavior.

How could this be? Our interpretation of these results centers on the importance of the relative expression levels within a population. In single populations, there are no relative expression levels because only one delta is being expressed. However, when more than one deltas are expressed in a cell, each surface expression level can be considered relative to the others also being expressed. We posit three possible explanations. In the first model, the relative proportion of co-expressed deltas influences *trans* binding by changing the probability that a delta on one cell will be given the opportunity to interact and bind with the same delta on another cell (Figure 3.19). For example, in a population that only expresses Pcdh7, the relative expression of that delta will always be 100% even if the absolute surface protein levels change because there is no other delta being expressed. In the experiments of Figure 3.16, single populations were tested against cells expressing two deltas in an approximate 1:1 ratio. Thus, the probability that a Pcdh7 molecule on the Pcdh7 expressing cells will interact with a Pcdh7 molecule in the Pcdh1+Pdh7 cells will always be 50%. Expressing more or less of Pcdh7 in the Pcdh7 only cells will not change that probability. Therefore, the outcome does not significantly change.

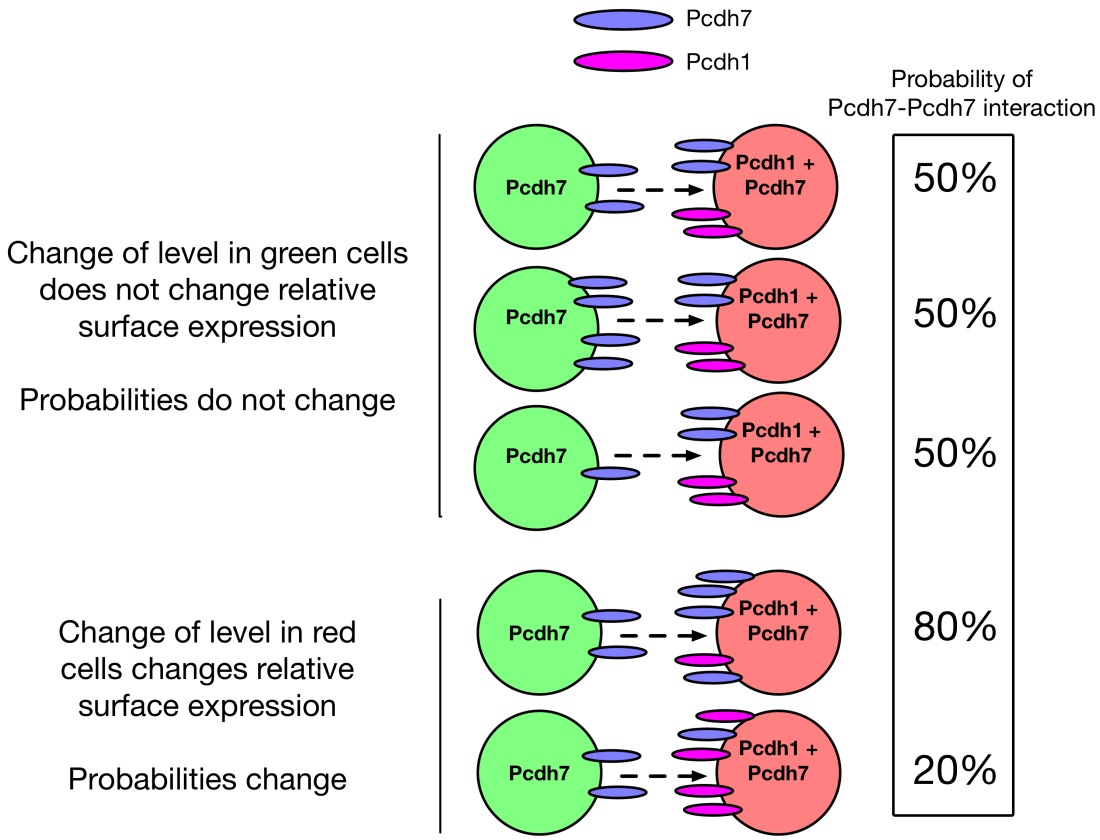

Figure 3.19 One possible model for how surface expression levels only effects cells expressing two deltas in mismatch coaggregation assay. In the top three cases, the surface level changes in the cells expressing only Pcdh7, but the probability that any one molecule of Pcdh7 on those populations will interact with a Pcdh7 molecule on the red cells is always 50%. The probability cannot change by altering levels until Pcdh7 is co-expressed with another delta (bottom two cells), and the amount at which they are expressed alters their relative proportions.

In contrast, if the proportion of co-expressed deltas are altered, dramatic changes are seen. For example, if Pcdh1+Pcdh7 are co-expressed in an approximate 1:1 ratio, a molecule of Pcdh1 on a cell just expressing Pcdh1 will have a 50% chance of interacting with a Pcdh1 molecule on the Pcdh1+Pcdh7 cells, and we know this leads to interfacing. On the other hand, a molecule of Pcdh7 on a cell just expressing Pcdh7 will also have a 50% chance of interacting with a Pcdh7 molecule on the Pcdh1+Pcdh7

cells, but this leads to intermixing. If we now change the proportion in the Pcdh1+Pcdh7 cells to such that Pcdh1 > Pcdh7, we see that the CoAgg index of Pcdh1 is greatly increased and that of Pcdh7 is greatly decreased. This may occur because the probability of Pcdh1-Pcdh1 *trans* interactions have now been increased and the probability of Pcdh7-Pcd7 interactions are decreased. Thus, the influence of Pcdh7's higher apparent affinity has been modulated by having a decreased surface expression. We tested a wide range of pairs across the hierarchy of apparent affinities and found that in every case, the CoAgg Index of a mismatch pair is widely influenced by the proportion of the co-expressed pair. The magnitude of change observed depends on how much the proportion was distorted, and also which deltas were being tested. Thus, despite any differences in affinity, the relative levels of each delta can "weigh" their influence on the adhesive nature between any two cell populations.

In a second possibility, differences in relative surface levels may alter how different deltas interact or are arranged within the same cell, and in turn this changes their adhesive binding function in *trans*. For example, differential or promiscuous *cis* binding may rearrange how the deltas are arrayed on the surface, and therefore influence their ability (or probability) to bind with the same delta on other cells. Or, *cis* interactions may differentially induce structural changes or block one delta from being able to bind in *trans*. A third and final possibility is that both cases may occur. Future work will need to be done to determine how surface levels influence coaggregation behaviors.

5.4. Comparison to Clustered Protocadherins

Our results are very different from the clustered protocadherins. First, none of the clustereds were able to intermix when two populations were mismatched, whereas here we show the deltas can produce diverse outcomes. Furthermore, the specific make up of clustered combinations did not matter - only whether or not the repertoire matched. In contrast, the specific repertoire matters for the deltas as they have different characteristics (apparent affinity). What could explain these differences? One possibility is that all the clustereds are much weaker in affinity than most of the deltas, and that none are strong enough to “overcome” the interference of the presence of a mismatched protocadherin. Thus, if the affinities of all the clustereds and deltas were combined to make a ranking, the clustereds may be primarily on the end with the weaker deltas (i.e. Pcdh19, Pcdh9, Pcdh18). This would also serve well for “self-avoidance”, where the adhesive states must disengage after isoneural homophilic recognition occurs.

We note that our experimental conditions differed in a few aspects from the clustered experiments. However, we believe our method considered and controlled for more areas of possible variation. First, we mixed cell populations an hour after electroporation, before gene expression turns on. This means that as the cells are beginning to express the deltas, each cell would have an equal opportunity to interact with both populations. In the prior experiments done by others with the clustered protocadherins, populations were mixed after recovering cells overnight. In our hands, this could lead to the formation of aggregates prior to being combined with another cell

population, and this could potentially lead to preference of homophilic aggregation. Secondly, our use of a controlled, slow rotational speed allowed us to detect weaker adhesive interactions between two populations. Whether weaker adhesive states such as interfacing exist for clustereds and were not detected due to inappropriate speeds remains to be determined. Third, our results consider and show the significance of surface expression levels. It remains to be seen how differences in surface level expression affect clustered protocadherin adhesion.

It is noteworthy that in evolution, the deltas emerged much earlier than the clustered protocadherins (Hulpiau et al. 2016; Ravi et al. 2016). The significance of what this means for functional differences has not been answered. One interpretation of cadherin evolution is that the "primordial" cadherins had many extracellular repeats, and as evolution progressed and adhesive interactions needed to be more specific the number of EC repeats decreased. We note that all the clustereds have 6 EC repeats while the δ-1s have 7 and the δ-2s have 6. One possibility is that the clustereds emerged from a duplication event of one of the weaker, lower affinity δ-2 protocadherins. However, sequence alignments do not reveal any obvious homology between a particular δ-2 and any of the clustereds. Since it is assumed that clustereds and deltas are co-expressed in neurons, an important step will be to incorporate the two sub-families to generate a more encompassing understanding of protocadherin function.

5.5. Proposed Model and *In Vivo* Relevance

We therefore propose a simple model to explain our results. The adhesive function of a combinatorial 'adhesive code' of δ-protocadherins depends on two factors that will dictate the adhesive interactions it will have with other populations. First, which deltas are expressed in a cell is critical because the apparent affinities among the deltas differ. Second, the influence of a given delta is weighted by the proportion of its relative surface expression to the other deltas that are also expressed within the cell (Figures 3.17, 3.18). Thus, we propose that while apparent affinity is fixed, the effective weight this has on the cell can be modulated by changing its relative amount on the cell surface. This can be achieved by expressing more or less of a given delta, or expressing multiple deltas (higher combinations) to "dilute" the presence of another. This dilution effectively reduces the impact of a high affinity delta, and enables the presence of weaker affinity deltas to exert an effect on cellular behavior.

How could this be useful for neurons to specify neural wiring choices? To date, the most established model in the protocadherin field is that of the clustered protocadherins. Through combinatorial expression, the ~58 clustered protocadherins have been theorized to produce an estimated 3×10^{10} unique surface combinations (Yagi 2012). However, except for a few instances (Chen et al. 2012), deletion studies require whole clusters to be deleted to see functional consequences (Wang et al. 2002; Weiner et al. 2005; Lefebvre et al. 2012; Hasegawa et al. 2017). This model suggests that in most cases, the availability of a large set of genes to create diversity is critical, rather than specific members. In contrast, genetic studies on the δ-protocadherins have shown that

knocking out or mutating a single member can have significant functional impacts on the nervous system (Uemura et al. 2007; Dibbens et al. 2008; Emond et al. 2009; Williams et al. 2011; Bruining et al. 2015). Additionally, individual members have been associated with neurological conditions such as autism and schizophrenia (Bucan et al. 2009; Hussman et al. 2011; Girirajan et al. 2013; Butler et al. 2015). This suggests that while diversity may be the general driver of clustered function in the nervous system, individual deltas have critical roles in neural function. In contrast to the clustered protocadherins, if a neuron has the potential to express all nine deltas but each can only be present in an "on" or "off" state (i.e. no variation in expression level), there are only 512 different combinatorial possibilities. However, we have shown that for the deltas, surface levels are a significant factor. By introducing a second state for the deltas, such that there are now three possible states (off, low, high), 19,683 different surface combinations can be generated. If there were 10 possible states, the number of unique surface identities now reaches 10^6. However, these calculations fail to take into account the impact of differential apparent affinities among the deltas. Notably, deltas of both low and high apparent affinities, as defined in this chapter, are known to have impacts on neural function. What does this mean for combinatorial identities and function? It suggests that the delta model becomes more complex because the meaning of the combinatorial repertoire of a cell depends on multiple factors.

Notably, deltas of both low and high apparent affinities, as defined in this chapter, have impacts on neural function. As the deltas appear to be less similar to the clustereds, this adds a degree of complexity in understanding their role in adhesion.

How is this relevant to *in vivo* processes? Because the deltas are expressed throughout the soma, axon, growth cone and synapses of neurons, their adhesive events may have wide functions. For example, it is now entirely conceivable that combinatorial codes of deltas could specify neural migrations, axon guidance patterns, and synaptogenesis. By progressing towards a comprehensive understanding of their function as a family, we can begin to interpret *in vivo* studies with more accuracy. For example, in knockout studies of protocadherins, we can now ask if the effect observed is due to the loss of that particular delta, or the ability of other deltas to have a stronger influence on the cell surface due to the absence of the delta that was removed. In neurological conditions such as Pcdh19 FLE, we may now begin to ask what the consequences are on the combinatorial repertoire of other deltas when Pcdh19 is mutated? The punctate and combinatorial expression of neurons may be one of the reasons why some conditions like Pcdh19 FLE are difficult to understand and model, but an approach driven by combinatorial methodologies such as those demonstrated here will likely provide means to reduce these complexities and begin to understand their functions *in vivo*.

We do note that there are unpublished reports (from our lab and others) that deltas can interact heterophilically in *cis*, and this is consistent with the clustereds which show promiscuous *cis* dimerization in the EC domain most proximal to the membrane. It is possible that *cis* complexes of deltas changes the nature of their *trans* binding. In addition, since several deltas have been shown to modulate other proteins (Emond et al. 2011; Chen & Gumbiner 2006; Yasuda et al. 2007), it is possible that interactions with other proteins on the cell surface may also alter their functions. Future studies will

be needed to explore how delta combinations may be affected through interactions with other proteins.

Even though the combinations tested here were low in complexity (i.e. two genes), we believe these results form a foundational framework for future studies to understand the functions of higher level δ-protocadherin combinations. In summary, we propose that despite the possibility of complex combinations, the adhesive outcome between two cell populations expressing δ-protocadherins depends on which deltas are expressed, their apparent affinities, and the relative proportions of their surface expression levels.

6. References

Bruining H, Matsui A, Oguro-Ando A, et al (2015) Genetic Mapping in Mice Reveals the Involvement of Pcdh9 in Long-Term Social and Object Recognition and Sensorimotor Development. Biological Psychiatry 78:485–495. doi: 10.1016/j.biopsych.2015.01.017

Bucan M, Abrahams BS, Wang K, et al (2009) Genome-Wide Analyses of Exonic Copy Number Variants in a Family-Based Study Point to Novel Autism Susceptibility Genes. PLOS Genet 5:e1000536. doi: 10.1371/journal.pgen.1000536

Butler M, Rafi S, Hossain W, et al (2015) Whole Exome Sequencing in Females with Autism Implicates Novel and Candidate Genes. International journal of molecular sciences 16:1312–1335. doi: 10.3390/ijms16011312

Chen WV, Alvarez FJ, Lefebvre JL, et al (2012) Functional Significance of Isoform Diversification in the Protocadherin Gamma Gene Cluster. Neuron 75:402–409. doi: 10.1016/j.neuron.2012.06.039

Cooper SR, Jontes JD, Sotomayor M, Weis WI (2016) Structural determinants of adhesion by Protocadherin-19 and implications for its role in epilepsy. eLife 5:e18529. doi: 10.7554/eLife.18529

Dibbens LM, Tarpey PS, Hynes K, et al (2008) X-linked protocadherin 19 mutations cause female-limited epilepsy and cognitive impairment. Nature genetics 40:776–781. doi: 10.1038/ng.149

Duguay D, Foty RA, and Steinberg MS (2003). Cadherin-mediated cell adhesion and tissue segregation: qualitative and quantitative determinants. Developmental Biology 253, 309–323.

Emond MR, Biswas S, Blevins CJ, and Jontes JD (2011). A complex of Protocadherin-19 and N-cadherin mediates a novel mechanism of cell adhesion. J. Cell Biol. 195, 1115–1121.

Emond MR, Biswas S, and Jontes JD (2009). Protocadherin-19 is essential for early steps in brain morphogenesis. Developmental Biology 334, 72–83.

Girirajan S, Johnson RL, Tassone F, et al (2013) Global increases in both common and rare copy number load associated with autism. Human Molecular Genetics 22:2870–2880. doi: 10.1093/hmg/ddt136

Goodman KM, Rubinstein R, Thu CA, et al (2016b) Structural Basis of Diverse Homophilic Recognition by Clustered α- and β-Protocadherins. Neuron 90:709–723. doi: 10.1016/j.neuron.2016.04.004

Goodman KM, Rubinstein R, Thu CA, et al (2016a) γ-Protocadherin structural diversity and functional implications. eLife 5:e20930. doi: 10.7554/eLife.20930

Hasegawa S, Kobayashi H, Kumagai M, et al (2017) Clustered Protocadherins Are Required for Building Functional Neural Circuits. Frontiers in molecular neuroscience 10:114. doi: 10.3389/fnmol.2017.00114

Hayashi S, Inoue Y, Kiyonari H, et al (2014) Protocadherin-17 mediates collective axon extension by recruiting actin regulator complexes to interaxonal contacts. Developmental cell 30:673–687. doi: 10.1016/j.devcel.2014.07.015

Hertel N, Krishna-K, Nuernberger M, Redies C (2008) A cadherin-based code for the divisions of the mouse basal ganglia. The Journal of comparative neurology 508:511–528. doi: 10.1002/cne.21696

Hirano S, Yan Q, Suzuki ST (1999) Expression of a novel protocadherin, OL-protocadherin, in a subset of functional systems of the developing mouse brain. Journal of Neuroscience 19:995–1005.

Hulpiau P, Gul IS, Van Roy F (2016) Evolution of Cadherins and Associated Catenins. The Cadherin Superfamily 13–37. doi: 10.1007/978-4-431-56033-3_2

Hussman JP, Chung R-H, Griswold AJ, et al (2011) A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. Molecular autism 2:1. doi: 10.1186/2040-2392-2-1

Izuta Y, Taira T, Asayama A, et al (2015) Protocadherin-9 involvement in retinal development in Xenopus laevis. Journal of biochemistry 157:235–249. doi: 10.1093/jb/mvu070

Krishna-K K, Hertel N, and Redies C (2011). Cadherin expression in the somatosensory cortex: evidence for a combinatorial molecular code at the single-cell level. Neuroscience 175, 37–48.

Lefebvre JL, Kostadinov D, Chen WV, et al (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature 1–7. doi: 10.1038/nature11305

Lin J, Wang C, Redies C (2012a) Expression of delta-protocadherins in the spinal cord of the chicken embryo. The Journal of comparative neurology 520:1509–1531. doi: 10.1002/cne.22808

Lin J, Yan X, Wang C, et al (2012b) Anatomical expression patterns of delta-protocadherins in developing chicken cochlea. Journal of Anatomy 221:598–608. doi: 10.1111/j.1469-7580.2012.01568.x

Nicoludis JM, Vogt BE, Green AG, et al (2016) Antiparallel protocadherin homodimers use distinct affinity- and specificity-mediating regions in cadherin repeats 1-4. eLife 5:213. doi: 10.7554/eLife.18449

Ravi V, Yu W-P, Pillai NE, et al (2016) Cyclostomes Lack Clustered Protocadherins. Molecular biology and evolution 33:311–315. doi: 10.1093/molbev/msv252

Redies C (2003) Cadherins as regulators for the emergence of neural nets from embryonic divisions. Journal of Physiology-Paris 97:5–15. doi: 10.1016/j.jphysparis.2003.10.002

Rubinstein R, Thu CA, Goodman KM, et al (2015) Molecular Logic of Neuronal Self-Recognition through Protocadherin Domain Interactions. Cell 163:1–15. doi: 10.1016/j.cell.2015.09.026

Sano K, Tanihara H, Heimark RL, Obata S, Davidson M, St John T, Taketani S and Suzuki S (1993). Protocadherins: a large family of cadherin-related molecules in central nervous system. Embo J. *12*, 2249–2256.

Schreiner D, Weiner JA (2010) Combinatorial homophilic interaction between gamma-protocadherin multimers greatly expands the molecular diversity of cell adhesion. Proceedings of the National Academy of Sciences of the United States of America 107:14893–14898. doi: 10.1073/pnas.1004526107

Shapiro L, and Colman DR (1999). The diversity of cadherins and implications for a synaptic adhesive code in the CNS. Neuron *23*, 427–430.

Sperry RW (1963). Chemoaffinity In The Orderly Growth Of Nerve Fiber Patterns And Connections. Proceedings of the National Academy of Sciences of the United States of America 50:703–710.

Tai K, Kubota M, Shiono K, Tokutsu H, and Suzuki S (2010). Adhesion properties and retinofugal expression of chicken protocadherin-19. Brain Res. 1344, 13–24.

Thu CA, Chen WV, Rubinstein R, et al (2014) Single-cell identity generated by combinatorial homophilic interactions between α, β, and γ protocadherins. Cell 158:1045–1059. doi: 10.1016/j.cell.2014.07.012

Uemura M, Nakao S, Suzuki ST, Takeichi M, and Hirano S (2007). OL-Protocadherin is essential for growth of striatal axons and thalamocortical projections. Nat. Neurosci. 10, 1151–1159.

Vanhalst K, Kools P, Staes K, van Roy F, and Redies C (2005). delta-Protocadherins: a gene family expressed differentially in the mouse brain. Cell. Mol. Life Sci. 62, 1247–1259.

Wang X, Weiner JA, Levi S, et al (2002) Gamma protocadherins are required for survival of spinal interneurons. Neuron 36:843–854.

Weiner JA, Wang X, Tapia JC, Sanes JR (2005) Gamma protocadherins are required for synaptic development in the spinal cord. Proceedings of the National Academy of Sciences of the United States of America 102:8–14. doi: 10.1073/pnas.0407931101

Williams EO, Sickles HM, Dooley AL, et al (2011) Delta Protocadherin 10 is Regulated by Activity in the Mouse Main Olfactory System. Frontiers in Neural Circuits. doi: 10.3389/fncir.2011.00009

Wöhrn JC, Nakagawa S, Ast M, et al (1999) Combinatorial expression of cadherins in the tectum and the sorting of neurites in the tectofugal pathways of the chicken embryo. Neuroscience 90:985–1000.

Yoshida K (2003) Fibroblast cell shape and adhesion in vitro is altered by overexpression of the 7a and 7b isoforms of protocadherin 7, but not the 7c isoform. Cellular & molecular biology letters 8:735–741.

Chapter 4 Discussion

1. Abstract

Understanding how the nervous system develops and becomes wired presents a major challenge in neuroscience. A widely accepted but poorly understood model for how neurons can generate complex wiring specificities with only a small number of genes is through the use of combinatorial expression patterns within individual neurons. However, current methods are not able to produce a model for understanding how these different genes may function together towards the goal of wiring the nervous system. In this thesis, the role of combinatorial expression and function of the δ-protocadherins were studied and defined. Utilizing single neuron expression analysis and tissue culture aggregation assays, these studies demonstrate how a small number of genes, when expressed in combination, can produce diverse outcomes. In this discussion, the possibilities and limitations of extending this combinatorial methodology to the remaining members of the cadherin superfamily to generate a larger understanding of combinatorial function is explored.

2. The Importance and Functional Implications of Combinatorial Gene Function

In Chapter 1, the limitations of transcriptome profiling and genetic analyses for understanding the combinatorial nature of neural wiring were described (Figure 4.1). When the transcriptome is profiled, thousands of genes can be defined, but very little understanding of how these genes function together is gained. In genetic studies on single genes, a detailed understanding of that gene's individual function is gained, but how that gene interacts with the many other genes also being expressed is not considered. For example, many different genes have been identified to play crucial roles in the process of neural wiring. However, current approaches are inefficient at capturing and understanding this complexity. To begin tackling complex questions involving many genes, we must move towards connecting these two disparate ends into an integrated model that can better explain what occurs *in vivo*. In an ideal case, experiments would provide high functional insight for a high number of genes. However, current strategies do not accommodate both, and usually as one aspect is gained, the other is lost. Considering this, how can new strategies be applied for enhancing our understandings of the complexities of nervous system development?

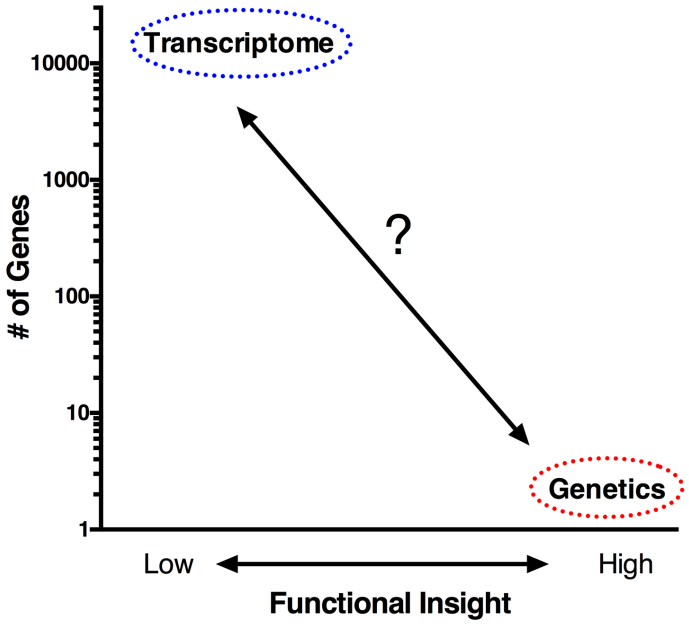

Figure 4.1 Limitations of traditional approaches in understand complex combinatorial gene expression and function.

3. Summary of Results and Working Model for δ-Protocadherin Adhesion

The aim of this thesis was to demonstrate a method for understanding combinatorial gene function in cell adhesion using a small gene family, the δ-protocadherins. Using single-cell analyses and *in vitro* assays, our results provide an enhanced framework for further studying and interpreting δ-protocadherin functions *in vivo* by revealing how they function in simple combinations. In Chapter 2, the extent of combinatorial expression of δ-protocadherins was defined in single olfactory sensory neurons, and we found that deltas are expressed in a wide range of combinations in single neurons. These expression patterns are consistent with the deltas generating combinatorial codes,

where individual neurons express different repertoires of deltas to provide different adhesive functions.

In Chapter 3, the adhesive functions of deltas alone, and in simple combinations were characterized. We concluded that within the family, the deltas possess different apparent affinities which create significant functional differences between them. This emphasizes that the specific identify of the combinatorial repertoire matters, as which deltas are expressed becomes important in dictating its adhesive function. Furthermore, we show that the relative proportion of each delta at the cell surface can act to modulate the differences defined by apparent affinity. Thus, the final adhesive outcomes of delta combinations ultimately depends on which deltas are expressed and their relative level of surface expression. Because deltas have been found to be expressed throughout the nervous system (Hirano et al. 1999; Gaitan and Bouchard 2006; Hertel et al. 2008; Redies et al. 2008; Krishna-K and Redies 2009; Krishna-K et al. 2011; Priddle and Crow 2013) and also on all parts of the neuron (Hirano et al. 1999a; Yasuda et al. 2007; Asahina et al. 2012), the impact of these combinatorial adhesive codes are likely to be widespread and critical for understanding several different aspects of neural development. As the deltas are part of a larger gene family, how can these results be utilized to expand to a more comprehensive model involving more genes?

4. Integrating a δ-Protocadherin Model with the Clustered Protocadherin Model

We know from ongoing studies that single olfactory sensory neurons co-express both deltas and clustered protocadherins (data not shown). Therefore, a simple and natural first step is to integrate our delta model with the clustered protocadherin model to create a larger, more encompassing model of combinatorial protocadherin function. As reviewed in the previous chapters, the clustered protocadherins consistent of ~58 different genes that are combinatorially expressed in most neurons. However, their model of function is different from our model for the deltas. In the clustered protocadherins, cell aggregation assays will always segregate if there is just one mismatched clustered is present between the two populations. Genetic studies have shown that in most neural systems, no one clustered is necessary, and effects are generally only observed by removing a whole cluster. In contrast, knocking out a single delta can have a drastic impact on neural development and wiring. Therefore, the immediate question to ask is how do these two families function together to influence cellular behavior, and under what model?

To do this, the approach outlined in this thesis can be repeated. Because it is important to know for aggregation assays the extent of co-expression, the combinatorial expression of deltas and clustered protocadherins must first be characterized in single neurons. Next, their adhesive functions can begin to be characterized. A critical component of our delta model is that they display different apparent affinities. Although

it appears that the clustereds may not vary as much in this respect based on the current model, it will be critical to compare them to produce a combined hierarchy to understand how they function when present in the same cell. However, the clustered model was based on experiments testing only the members that had the highest sequence identities (greater than 80%) in order to test the most stringent adhesive conditions for specificity (Thu et al. 2014). The authors posited that if the clustereds with the highest sequence similarities failed to generate heterophilic binding, then the ones with lower sequence similarities would be highly unlikely to do so either. However, one possibility is that this selection caused them to sample clustereds with all very similar affinities, and therefore missed the role of divergent sequences may have on generating different adhesive coaggregation behaviors between populations expressing mismatched protocadherins.

To address this initially without having to test all 58 clustereds, select members representing the range of sequence identities within a given clusters (alpha, beta, and gamma) could be tested. This way, any behaviors generated by different affinities will not go overlooked. To do this, the various experiments using K562 cells in Chapter 3 can establish how the clustereds compare to the deltas, allowing them to be merged into our delta hierarchy of apparent affinities. For example, what happens to high affinity deltas (e.g. Pcdh7) when the clustered acts as a mismatch, and vice versa? Critically, this will attempt to produce a larger, integrated model explaining the combinatorial functions of a larger set of genes while maintaining high functional understanding.

Combining these two cadherin subfamilies will produce an integrated model for nearly all protocadherins.

However, there are a few other remaining members in the protocadherin family. Based on their sequences, they show low sequence identities to the clustereds and deltas (Figure 4.2), suggesting that they may provide additional complexity in combination through more affinity differences when multiple protocadherins are expressed. These remaining members could easily be fully integrated with the deltas and clustereds to provide a comprehensive model of all protocadherins. In Chapter 3, the combinatorial complexity tested was low (i.e. co-expressing only two protocadherins) to establish the foundations of the methodology. To address the amount of complexity observed *in vivo*, combinations could easily be extended to test more than two to demonstrate and define combinatorial function at greater complexities.

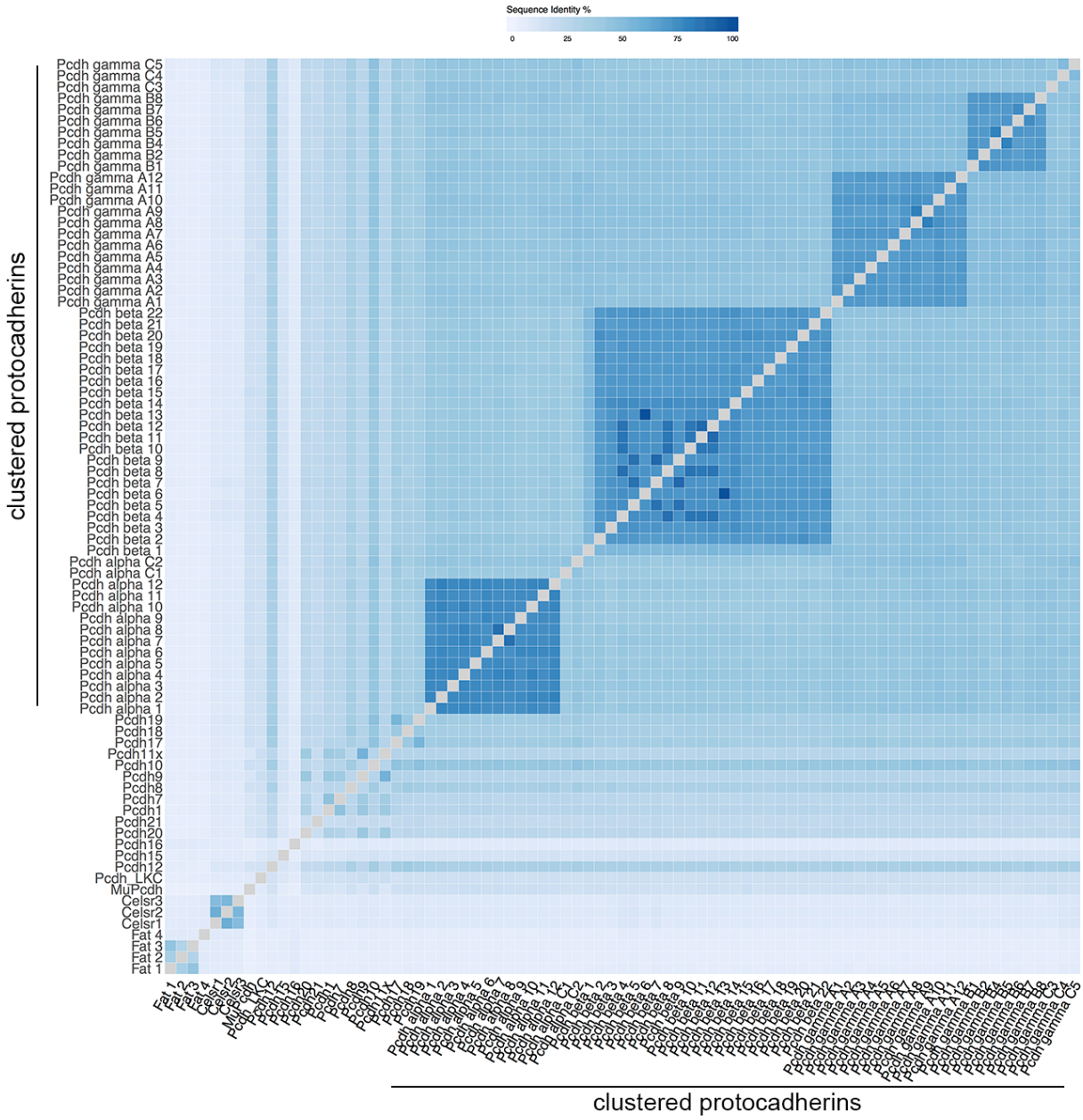

Figure 4.2 Heat map of all protocadherin extracellular domain sequences. Extracellular regions were annotated using prediction servers to mark the end of the signal peptide and the start of the transmembrane domain, and then the protein sequences were aligned using MUSCLE to determine the percent sequence identity between all members. Note that the blocks of high homology (large blue squares) represent the alpha, beta, and gamma clustered protocadherins.

5. Expanding Towards an Integrated Model of Combinatorial Cadherin Adhesion

As noted in Chapter 1, the cadherins are the largest group of cell adhesion molecules. While we have focused on one subfamily, the δ-protocadherins, classical cadherins such as N-cadherin have been extensively studied. Critically, some interactions are known to occur among different cadherin subfamily members that can alter their function. For example, Pcdh19 and N-cadherin have been shown to form a complex which generates a novel adhesive specificity that no longer recognizes Pcdh19 or N-cadherin alone (Biswas et al. 2010). In contrast, when the clustered protocadherins are co-expressed with N-cadherin, N-cadherin adhesion appears to be dominant and the clustereds appear to not contribute towards adhesive specificity (Thu et al. 2014). Therefore, it suggests that there are underlying hierarchies and rules for how cadherin adhesion occurs in combinations, and that unless they are studied and characterized as such, their true functions will never be understood.

Our ongoing single-cell studies suggest that individual neurons can express many different members of the cadherin superfamily and thus exhibit a high level of combinatorial complexity, yet almost nothing is known about how members of the cadherin superfamily function together when expressed in combinations. Could the approach outlined above be expanded further to include all cadherins, to produce an integrated model for combinatorial adhesion among all cadherins? Conceivably, the same approach could be used produce a comprehensive integrated model of the ~110

members of the cadherin superfamily (Hulpiau et al. 2016). A functional understanding of how this large gene family functions together would be immensely powerful in understanding how the cadherin superfamily controls many aspects of neural development and wiring.

Since the cadherin superfamily consists of over 100 genes, they cannot be easily tested at once. How could integration be accomplished? One method to tackle this problem is to utilize a modular approach towards constructing larger models of combinatorial gene function. In this approach, each sub family can be studied to produce a module which represents the combinatorial function for that specific group. Next, modules can be combined into larger modules encompassing more genes until a singular model is constructed which explains and predicts the adhesive outcomes of combinations consisting of all cadherins.

This problem can be established within the framework of the concept outlined in Figure 4.1. For example, where expression studies can inform us about which cadherin genes are expressed in single cells or tissues, there is no functional insight gained into their combined function (Figure 4.3A). In contrast, genetic studies on single cadherins provide functional insight into that individual cadherin, but do not inform how they work with the other cadherins that are also present in the same cells. The current models established in chapters 2 and 3 for the deltas, as well as the one established for the clustered protocadherins by others represent modules of sub-families that increase our functional insight into a by providing meaning for how function in combination. However, they remain distinct and separate until they are combined, as described above (Figure

4.3B). For example, this can be done by combining clustered and δ-protocadherins together in the same cell, to see which subfamily exerts a greater effect on aggregation. This would produce a larger module that considers more genes but does not compromise functional description and insight (Figure 4.3C). Finally, after this process is repeated several times to create modules for each sub family, they can each be integrated into a larger, single module consisting of all members of the cadherin superfamily (Figure 4.3D). Notably, this approach builds vertically on the graph, which maintains high functional insight despite consisting of more genes. Thus, this approach provides a new way to overcome the limitations of current approaches where this would not be possible.

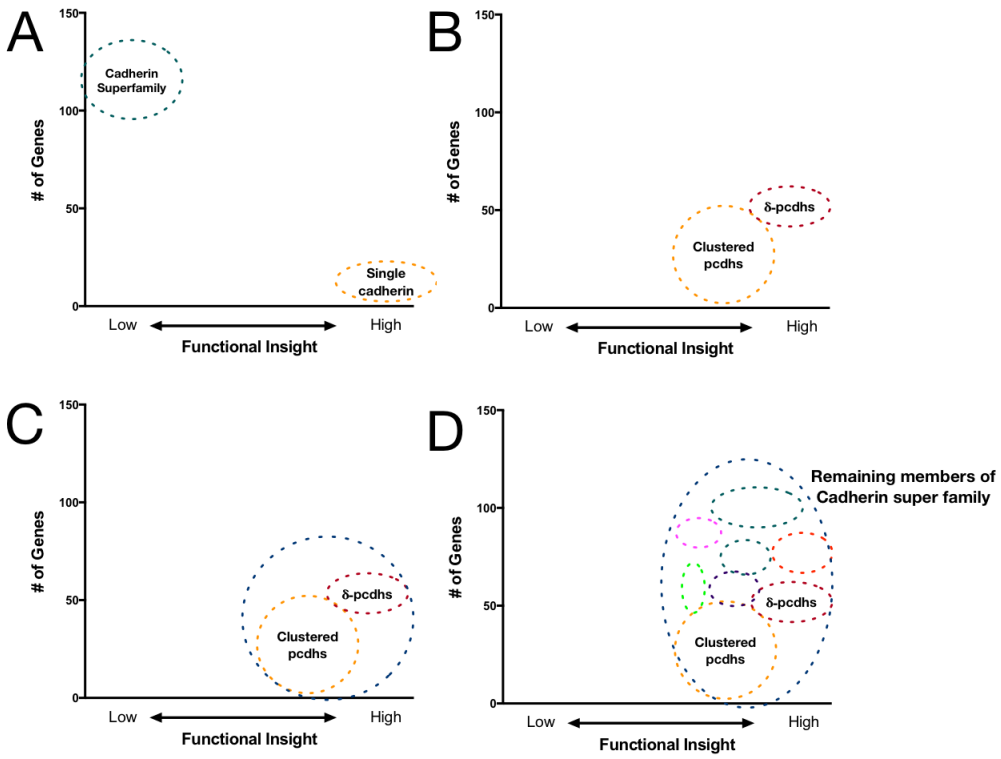

Figure 4.3 A combinatorial methodology towards understanding combinatorial cadherin expression and function through modular experiments.

6. Possible Limitations and *In Vivo* Relevance

There are several caveats to this approach to consider. First, these experiments rely on a non-neural system (K562 cells) to provide an understanding of function in the nervous system. However, compared to neuronal cultures which exhibit high levels of heterogeneity, tissue culture systems such as K562 are a homogenous population, and do not endogenously express any cadherins. Therefore, they provide an "empty vessel" to express only specific combinations of cadherins, thereby simplifying the interpretation of adhesive events.

Second, most cadherins have intracellular domains which participate in a variety of signaling processes, although many are poorly characterized. In the experiments in Chapter 3, the ICDs were removed such that only the extracellular and transmembrane domains were expressed. This allows the experiment to measure only the adhesive interactions, and not signaling events that may occur downstream. However, if the ICDs were used, signaling events are likely to not be appropriately measured in K562 cells as they are non-neuronal and therefore likely to have different signaling machinery. Furthermore, signaling has been shown to be dependent on where in the nervous system you are studying. Thus, signaling events modeled in this system would not be useful toward generalizing towards the whole nervous system. For these reasons, the lack of ICD domains does not pose a major limitation for these experiments. Instead, this system defines the rules for the first step - the adhesive interactions. However, signaling events remain critical to understand and would need to be further studied in

other experiments. For example, some cadherins have been shown to modulate each other (Chen and Gumbiner 2006; Yasuda et al. 2007), and the ability of several different axon guidance cues to cross-talk to produce additive (Taku et al. 2016) (Funato et al. 2000), hierarchical (Stein 2001; Chang et al. 2004), and modulatory or synergistic outcomes (Stevens & Jacobs 2002; Höpker et al. 1999; Rhee et al. 2002; Bonanomi et al. 2012; Bielle et al. 2011) further demonstrates the importance of studying genes in combinatorial as these functions would not be detected if studied alone. Defining the possibilities of combinatorial extracellular binding provides the first step for future studies to define the signaling events these binding events may initiate.

Another difficulty in interpreting *in vitro* studies for *in vivo* function is that the nervous system is very dynamic. For example, some attractive guidance molecules can also act as repulsive cues, depending on the developmental stage, location, and cell type under consideration (Colamarino and Tessier-Lavigne 1995; Wong et al. 1997). The expression, and response of a neuron to certain cues can change depending on where the neuron is - even as the axon extends from one area towards another. Thus, functional interpretations of combinations may not always apply *in vivo*. For example, the impact of clustered protocadherins has been shown to vary depending upon neuron type and location in the nervous system (Lefebvre et al. 2012; Molumby et al. 2016; Chen et al. 2017; Mountoufaris et al. 2017). Therefore, combinatorial functions will need to be carefully interpreted depending on which neural system is being studied.

Another aspect that needs to be considered is that the combinatorial repertoire defined in a cell does not mean that the entire repertoire is present in an equal distribution

throughout the whole neuron. Because neurons are known to have sub-cellular localization of mRNAs (Brittis et al. 2002), it is possible that even though a neuron may express a given combination, only a subset of that repertoire may be expressed in the growth cone, the soma, or the axon, for example. As single-cell technologies improve, these sub-cellular localizations will be better defined and allow for more precise spatial interpretations of these combinations.

Despite these limitations, experiments could be performed to better interpret cadherin functions *in vivo*. For example, one possibility would be to use sub-populations of neurons that are functionally the same. If distinct sub-populations express similar cadherin combinations but different from other sub-populations, it would suggest that the given combination is providing adhesive instructions. Sub-populations could include odorant receptor sub-types in olfactory sensory neurons which target specific glomeruli in the olfactory bulb, or neurons in specific cortical layers which may use combinatorial profiles to specify laminar migration and/or synapse specificities between layers. With CRISPR technology now widely available, mouse models can be easily designed to test combinatorial functions in defined circuitries to confirm and validate *in vivo* function.

7. Conclusion

This thesis outlines a possible approach to expand our knowledge of protocadherins and cadherin adhesive functions in combination. With the use of single-cell technologies to define combinations, and cell aggregation assays to understand the functional

implications of these combinations, we can begin to unravel the complex processes in neural wiring to better understand neural development.

8. References

Asahina H, Masuba A, Hirano S, Yuri K (2012) Distribution of protocadherin 9 protein in the developing mouse nervous system. Journal of Neuroscience 225:88–104. doi: 10.1016/j.neuroscience.2012.09.006

Biswas S, Emond MR, Jontes JD (2010) Protocadherin-19 and N-cadherin interact to control cell movements during anterior neurulation. The Journal of cell biology 191:1029–1041. doi: 10.1083/jcb.201007008

Brittis PA, Lu Q, Flanagan JG (2002) Axonal protein synthesis provides a mechanism for localized regulation at an intermediate target. Cell 110:223–235.

Chen WV, Nwakeze CL, Denny CA, et al (2017) Pcdhαc2 is required for axonal tiling and assembly of serotonergic circuitries in mice. Science 356:406–411. doi: 10.1126/science.aal3231

Chen X, Gumbiner BM (2006) Paraxial protocadherin mediates cell sorting and tissue morphogenesis by regulating C-cadherin adhesion activity. The Journal of cell biology 174:301–313. doi: 10.1083/jcb.200602062

Colamarino SA, Tessier-Lavigne M (1995) The axonal chemoattractant netrin-1 is also a chemorepellent for trochlear motor axons. Cell 81:621–629.

Gaitan Y, Bouchard M (2006) Expression of the delta-protocadherin gene Pcdh19 in the developing mouse embryo. Gene expression patterns : GEP 6:893–899. doi: 10.1016/j.modgep.2006.03.001

Hertel N, Krishna-K, Nuernberger M, Redies C (2008) A cadherin-based code for the divisions of the mouse basal ganglia. The Journal of comparative neurology 508:511–528. doi: 10.1002/cne.21696

Hirano S, Ono T, Yan Q, et al (1999a) Protocadherin 2C: A New Member of the Protocadherin 2 Subfamily Expressed in a Redundant Manner with OL-Protocadherin in the Developing Brain. Biochemical and Biophysical Research Communications 260:641–645. doi: 10.1006/bbrc.1999.0950

Hirano S, Yan Q, Suzuki ST (1999b) Expression of a novel protocadherin, OL-protocadherin, in a subset of functional systems of the developing mouse brain. Journal of Neuroscience 19:995–1005.

Hulpiau P, Gul IS, and Van Roy F (2016). Evolution of Cadherins and Associated Catenins. In The Cadherin Superfamily: Key Regulators of Animal Development and Physiology, S.T. Suzuki, and S. Hirano, eds. (Tokyo: Springer Japan), pp. 13–37.

Krishna-K K, Hertel N, and Redies C (2011). Cadherin expression in the somatosensory cortex: evidence for a combinatorial molecular code at the single-cell level. Neuroscience 175, 37–48.

Krishna-K, Redies C (2008) Expression of cadherin superfamily genes in brain vascular development. Journal of Cerebral Blood Flow 38; Metabolism 29:224–229. doi: 10.1038/jcbfm.2008.123

Lefebvre JL, Kostadinov D, Chen WV, et al (2012) Protocadherins mediate dendritic self-avoidance in the mammalian nervous system. Nature 1–7. doi: 10.1038/nature11305

Molumby MJ, Keeler AB, Weiner JA (2016) Homophilic Protocadherin Cell-Cell Interactions Promote Dendrite Complexity. Cell Reports 15:1037–1050. doi: 10.1016/j.celrep.2016.03.093

Mountoufaris G, Chen WV, Hirabayashi Y, et al (2017) Multicluster Pcdh diversity is required for mouse olfactory neural circuit assembly. Science 356:411–414. doi: 10.1126/science.aai8801

Priddle TH, Crow TJ (2013) Protocadherin 11X/Y a human-specific gene pair: an immunohistochemical survey of fetal and adult brains. Cerebral cortex 23:1933–1941. doi: 10.1093/cercor/bhs181

Redies C, Heyder J, Kohoutek T, et al (2008) Expression of protocadherin-1 (Pcdh1) during mouse development. Developmental dynamics 237:2496–2505. doi: 10.1002/dvdy.21650

Thu CA, Chen WV, Rubinstein R, et al (2014) Single-cell identity generated by combinatorial homophilic interactions between α, β, and γ protocadherins. Cell 158:1045–1059. doi: 10.1016/j.cell.2014.07.012

Wong JT, Yu WT, O'Connor TP (1997) Transmembrane grasshopper Semaphorin I promotes axon outgrowth in vivo. Development 124:3597–3607.

Yasuda S, Tanaka H, Sugiura H, et al (2007) Activity-Induced Protocadherin Arcadlin Regulates Dendritic Spine Number by Triggering N-Cadherin Endocytosis via TAO2β and p38 MAP Kinases. Neuron 56:456–471. doi: 10.1016/j.neuron.2007.08.020